The influence of First Language on playing brass instruments: An ultrasound study of Tongan and New Zealand trombonists

By Matthias Heyne

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Linguistics

University of Canterbury

2016

Table of contents

		page
List of figures		**1**
List of tables		**4**
Abbreviations and system for specifying pitch		**6**
Acknowledgements		**7**
Vita		**10**
Abstract		**11**
1	**Introduction**	**12**
2	**The Acoustics and pedagogy of brass instruments and previous research**	**16**
2.1	**Brass instrument acoustics: A simplified model**	**16**
2.2	**Further considerations**	**21**
2.3	**Vocal tract influence on brass instrument sound**	**24**
2.3.1	***In vivo* measurement of vocal tract influence on brass instrument sound**	**28**
2.3.1.1	**Direct measurement of vocal tract influence on brass instrument sound**	**28**
2.3.1.2	**Indirect measurement of vocal tract influence on brass instrument sound**	**32**
2.3.2	**The role of the glottis**	**34**
2.3.3	**Summary**	**36**
2.4	**Brass pedagogy**	**37**
2.4.1	**The use of speech syllables in brass pedagogy**	**37**
2.4.1.1	**Pedagogical writings on brass playing predating the modern era**	**37**
2.4.1.2	**Pedagogical writing on brass playing published in the last fifty years**	**43**
2.4.1.3	**Summary of speech syllables used in brass pedagogy**	**48**
2.4.2	**A personal account of recent developments in the world of brass playing**	**49**

		page
2.5	**Previous empirical research on brass playing**	**55**
2.5.1	**X-ray studies 1954-1975**	**55**
2.5.2	**Observations of the brass player’s lips**	**58**
2.5.3	**Magnetic Resonance Imaging**	**59**
2.5.4	**Ultrasound imaging of the tongue**	**64**
2.5.5	**Electromagnetic articulography**	**66**
2.5.6	**Conclusion**	**67**
2.6	**Perception of (brass) instrument sound by expert and non-expert listeners**	**68**
2.7	**Previous attempts at investigating language influence on brass playing**	**73**
2.8	**Chapter summary**	**76**
3	**Physiology and motor control of the upper vocal tract during speech production and brass playing**	**77**
3.1	**Upper vocal tract physiology and articulator movement**	**77**
3.1.1	**The jaw**	**78**
3.1.2	**The role of the jaw during speech production**	**79**
3.1.3	**The tongue**	**80**
3.2	**Human motor control**	**86**
3.2.1	**Models of human motor control**	**86**
3.2.2	**Generalized motor programs**	**88**
3.2.3	**Modular accounts of motor control**	**89**
3.2.4	**Is motor control optimal?**	**90**
3.3	**Motor control during speech production**	**91**
3.3.1	**Models of speech motor control**	**93**
3.3.2	**Articulatory setting theory**	**95**
3.3.3	**Modular accounts of speech motor control**	**96**
3.3.4	**Individual differences in vocal tract morphology and biomechanics and their influence on speech production**	**97**
3.4	**Control of the facial-oral-laryngeal-respiratory musculature outside speech production**	**102**

		page
3.5	Studies of motor behavior during brass playing	103
3.6	Similarities and differences between speech production and brass playing	106
3.6.1	The roles of the jaw and tongue	106
3.6.2	Is motor control shared across speech production and brass playing?	108
3.7	Predictions arising from a modular theory of motor control	109
3.8	Chapter summary	111
4	Possible areas of language influence on brass playing	112
4.1	Articulation: Beginning and connecting or ending notes	112
4.2	Tongue position during sustained note production	113
4.3	Hypotheses	113
5	Preliminary investigation: Online questionnaire	115
6	Methodology: Ultrasound imaging of the tongue (UTI)	121
7	Data collection	123
7.1	Selection of instrument group to record	123
7.2	Selection of participant languages	124
7.3	Brief overview of NZE and Tongan phonetics and phonology	125
7.3.1	NZE	125
7.3.2	Tongan	128
7.4	Speech elicitation: Wordlists	129
7.4.1	NZE wordlist	129
7.4.2	Tongan wordlist	130
7.5	Musical passages	131
7.6	Recording locations	131
7.7	The participants	132
7.7.1	NZE participants	132
7.7.2	Tongan participants	133

		page
7.8	**Recording setup**	**135**
7.8.1	**Ultrasound machine**	**135**
7.8.2	**Using UTI to record trombone players**	**135**
7.8.3	**Microphone placement for audio recordings**	**135**
7.9	**Recording procedure**	**137**
8	**Data analysis**	**140**
8.1	**Video transcoding**	**140**
8.2	**Audio-video alignment correction**	**140**
8.3	**Speech segmentation**	**141**
8.3.1	**Speech segmentation using the HTK toolkit implemented in LaBB-CAT**	**141**
8.3.2	**Manual segmentation for early data**	**142**
8.4	**Annotation of musical passages**	**142**
8.5	**Selection of the temporal location for ultrasound images to be analyzed**	**142**
8.6	**Tongue contour tracing using GetContours in MATLAB**	**143**
8.7	**Outlier removal, data transformations and export from MATLAB**	**143**
8.7.1	**Outlier removal**	**143**
8.7.2	**Estimation of ‘virtual origin’ and transformation of data to polar coordinates**	**145**
8.7.3	**Cutting traces to avoid edge effects on average curves**	**147**
8.8	**Calculation of SSANOVA average curves in R**	**149**
8.9	**Z-scoring of data**	**151**
8.10	**Chapter summary**	**154**
9	**Results**	**155**
9.1	**Speech data**	**155**
9.2	**Music data**	**158**
9.3	**Comparison of the two language groups**	**160**

		page
9.4	**Quantification of differences between tongue contours for vowel and trombone note productions**	**165**
9.5	**Results for individual participants**	**170**
9.5.1	**NZE players who use a centralized tongue position during sustained note production**	**171**
9.5.2	**Tongan players who use a back vowel tongue position during sustained note production**	**172**
9.5.3	**Participants displaying a pattern more typical for the opposite language group**	**173**
9.5.4	**Consonant production and trombone articulation data for two participants**	**174**
9.6	**Chapter summary**	**178**
10	**Discussion**	**179**
10.1	**Evaluation of hypotheses**	**179**
10.2	**Group-level findings versus individual variability**	**180**
10.3	**Tongue position during brass playing as local optimization**	**181**
10.4	**Requirements of airflow during brass playing**	**182**
10.4.1	**Overall midsagittal tongue shape**	**182**
10.4.2	**Air channeling within the oral cavity**	**184**
10.4.3	**‘Laryngeal states’ during brass playing**	**185**
10.5	**Acoustical considerations related to vocal tract influence on brass instrument sound**	**186**
10.6	**Motor efficiency considerations affecting the position of the front of the tongue**	**189**
10.7	**Language influence**	**189**
10.8	**Further considerations**	**190**
10.9	**Summary**	**191**
10.10	**Confounds**	**192**
10.10.1	**National schools of playing and different playing styles**	**192**
10.10.2	**Heterogeneity of the two language groups**	**193**
10.10.3	**Ultrasound probe stabilization**	**194**

		page
10.10.4	Other Confounds	194
10.11	Implications	197
10.11.1	Implications for modular theories of motor control	197
10.11.2	Implications for brass pedagogy	197
10.12	Future directions	199
11	Conclusion	200
References		202
Appendices		241
A	Online questionnaire and corresponding Human Ethics approval	241
B	Materials relating to the ultrasound study, including Human Ethics approval	249
C	Z-scoring ratios and individual plots for all participants	270
D	Permissions from publishers	290

List of Figures

Figure	*shortened caption*	*page*
0.1	Alternative methods for specifying pitch.	6
2.1a	A simplified model of a lip-reed instrument.	17
2.1b	'Water Trumpet' using a water-filled channel to depict wave oscillations within lip-reed instruments.	17
2.2	Waveforms (left) and harmonic spectra (right) of the pressure variations in a trombone mouthpiece during the playing of four notes.	19
2.3	Impedance spectrum for a Bb bass trombone (slide in first position, valve not depressed).	19
2.4	The relationship between pitch and frequency.	21
2.5	"Standing wave patterns for the lowest four resonant modes of a Bb tenor trombone. …"	22
2.6	"A simplified schematic (not to scale) showing most of the elements controlled by the player, beginning with the pressure of the air in the lungs. …"	24
2.7	"Schematic figures show idealisations of the duct-valve interactions in the voice (a), a lip valve instrument (b) and a reed instrument (c). …"	26
2.8	"The playing frequency of an artificial trombone playing system as the slide is extended from the closed position (0 mm). …"	27
2.9	"In this semilog plot, Z_{Bore} [instrument impedance] and Z_{Mouth} [vocal tract impedance] are plotted for the notes written (A) C5 and (B) G6. In both cases, the note is played and Z_{Bore} is measured with no valves depressed. …"	29
2.10	"Frequency of the first two maxima in the vocal tract impedance Z_{mouth} (resonances) compared with the playing frequency and the next two harmonics. …"	31
2.11	The impedance spectra of a vocal tract measured at the lips: glottis closed (red) and exhaling (black) …	36
3.1	Schematic of the modeled muscle groups and their attachments to the jaw and hyoid bone.	79
3.2	"Tongue muscles: coronal cross-section through the mid-tongue; the location of the cross-section is indicated by the vertical line through the sagittal tongue image at the bottom. …"	82

Figure	*shortened caption*	*page*
3.3	“Extrinsic tongue muscles: right side view. Geniohyoid and mylohyoid are included for the context only...”	83
3.4	“The three largest modes of variation in hard palate shape, determined in completely data-driven fashion, without imposing any prior notions about expected shape variations, by applying PCA to the observed hard palate shapes from the subject pool. …”	99
5.1	Ordered plot of the coefficient values for the various categories of ‘schools learned’ as estimated by the best fit-model on language influence reported in table 5.2 above.	119
7.1	The New Zealand English short front vowel shift. …	126
7.2	The monophthongs of New Zealand English. …	127
7.3	Vowel plots for primary stress versus unstressed vowels. F1 x F2 clouds show one standard deviation from mean value.	129
8.1	Tongue and palate traces for participant S12 NZE prior to (top plot) and after correcting for unwanted ultrasound transducer movement (bottom plot).	145
8.2	Estimation of the ‘virtual origin’ and pixel scale from a randomly selected ultrasound image by overlaying various lines.	146
8.3	Image illustrating the cutting off of extreme values along fan lines extending from the virtual origin for S29 NZE…	147
8.4	SSANOVA average curves with error bounds (red) and underlying traces (black) for selected vowels and notes produced by S29 NZE...	151
8.5	Plots illustrating the effect of the z-scoring procedure on the SSANOVA average curves for monophthong productions by three Tongan (top row) and three NZE participants (bottom row). …	154
9.1	SSANOVA average curves for the z-scored tokens of the five vowels of Tongan …, produced in accented position and averaged across the articulations of ten participants. …	156
9.2	SSANOVA average curves for the z-scored tokens of the stressed monophthongs of NZE plus schwa in non-final and final position, averaged across the productions of nine participants. …	157
9.3	SSANOVA average curves for the five vowels of Tongan and five different sustained notes played on the trombone by the same speakers…	159

Figure	*shortened caption*	*page*
9.4	SSANOVA average curves for the monophthongs of NZE and five different sustained notes played on the trombone by the same speakers…	160
9.5	Average tongue curves for sustained note productions by trombone players from the two different language groups ...	162
9.6	Matching (top row) and mismatched plots (bottom row) for the two language groups’ vowel productions, overlaid with sustained note productions by speakers of the same (matching) or opposite (mismatched) language group.	163
9.7	Average tongue contours for selected monophthong productions by participants from the two different language groups (Tongan = solid lines, NZE = dashed lines); only vowels were selected that can be expected to be roughly similar in terms of their acoustics across the two languages.	165
9.8	Schematic illustration of the area difference measurements implemented in MATLAB, forming the basis for the numerical comparisons in the following figures.	166
9.9	Average tongue contours for the five vowels of Tongan, produced by, and normalized across the ten Tongan participants of this study, overlaid on an image taken from figure 3: Five tongue segments of the Fisher-Logemann dataset; reproduced with permission from Stone, Epstein and Iskarous, 2004, p. 511.	167
9.10a	Line plots reporting the measured area in between average tongue contours for vowels (x-axis) and the five different notes (y-axis) for the back of the tongue. …	169
9.10b	Line plots reporting the measured area in between average tongue contours for vowels (x-axis) and the five different notes (y-axis) for the front of the tongue. …	170
9.11	Plots for 4 NZE participants whose average midsagittal tongue curves for sustained note production on the trombone pattern closely with or within the vicinity of one of the centralized vowels of NZE. …	172
9.12	Plots for two Tongan participants whose average midsagittal tongue curves for sustained note production on the trombone pattern closely with or within the vicinity of one of the back vowels of Tongan. …	173

Figure	*shortened caption*	*page*
9.13	Plots for two NZE and two Tongan participants whose average midsagittal tongue curves for sustained note production on the trombone assume a position that fits more closely with the pattern more typically displayed by players from the opposite language group. …	174
9.14	Average tongue contours for initial dental /t/s and interspeech postures (ISPs) produced by S4 Tongan, overlaid on monophthong productions by the same speaker (faint lines; left panel). Comparison of average tongue positions assumed during sustained note production in different registers of the trombone, coronal articulations produced during trombone playing and inter-playing postures (IPPs), overlaid on initial consonant productions and ISPs by the same participant (right panel).	176
9.15	Average tongue contours for initial alveolar /t/s and /d/s and interspeech postures (ISPs) produced by S5 NZE, overlaid on monophthong productions by the same speaker (faint lines; left panel). Comparison of average tongue positions assumed during sustained note production in different registers of the trombone, coronal articulations produced during trombone playing and inter-playing postures (IPPs), overlaid on initial consonant productions and ISPs by the same participant (right panel).	176
10.1	Revised vowel chart showing the division and overlap of articulatory regions.	183
10.2	Identical to figure 2.11.	187

List of Tables

Table	*shortened caption*	*page*
2.1a	Representative sample of articulation syllables used in early Italian sources.	40
2.1b	Representative sample of articulation syllables used in early French sources.	41
2.1c	Representative sample of articulation syllables used in early German sources.	42
2.2	National styles of brass playing.	54
5.1	Selected demographic and response data from the online questionnaire.	116

Table	*shortened caption*	*page*
5.2	Model outputs for the cumulative link model fit in R to predict the response to the question (1) Do you think that a person’s first language and/or other acquired languages have some kind of influence upon playing brass instruments?	118
5.3	Significance table for the various categories of ‘schools learned,’ determined by re-levelling the various factors.	119
5.4	Model outputs for the cumulative link model fit in R to predict the response to the question (2) “Which factor do you think is more influential in affecting brass playing, one’s First Language/s (and possibly Second Languages) or playing styles (nationals schools etc.)?”	120
7.1	Tongan consonant inventory (top) and vowel inventory (bottom).	128
7.2	Demographic data for the NZE participants included in this thesis.	133
7.3	Demographic data for the Tongan participants included in this thesis.	134
8.1a	Final numbers of vowel tokens for all NZE participants.	148
8.1b	Final numbers of sustained note tokens for all NZE participants.	148
8.2a	Final numbers of vowel tokens for all Tongan participants.	149
8.2b	Final numbers of sustained note tokens for all Tongan participants.	149
9.1	Token numbers for initial coronal consonants, ISPs, coronal articulations, and IPPs produced by S4 and S5.	177
10.1	List of confounds for this study	195-196

Abbreviations and system for specifying pitch

AE	American English
CNS	central nervous system
L1	First Language (native language)
L2	Second Language
NZ	New Zealand
NZE	New Zealand English
NZILBB	New Zealand Institute of Language Brain and Behaviour
SE	standard error
SSANOVA	smoothing spline analysis of variance
UTI	Ultrasound Tongue Imaging

Throughout this thesis, the US standard system will be used to specify pitch. Figure 0.1 below shows how these specifications translate to the Helmholtz system.

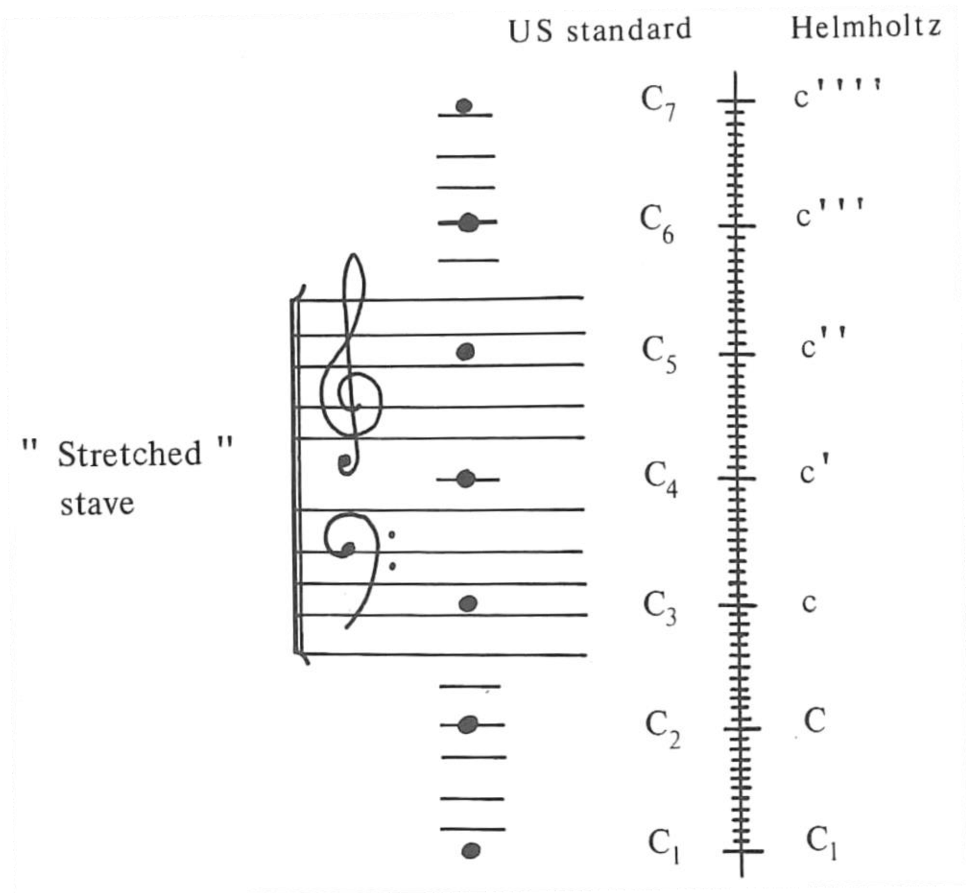


Figure 0.1: Alternative methods for specifying pitch. Reproduced with permission from Campbell & Greated, 1994, p. 73.

Acknowledgements

This thesis requires acknowledgements not only to all the wonderful people who helped me carry out my research at the University of Canterbury, and all the willing participants who supplied an hour's worth of data without compensation (most seemed to enjoy the experience!). It is also the product of interactions with a great number of fellow brass players who prompted me to first come up with an early version of my research question. After spending a year as an international exchange student in the USA while in High School, I studied orchestral and later jazz trombone in Germany, which allowed me to hear many colleagues from different language backgrounds play their instruments. It was at this point that I first started to notice that they all sounded a bit different. Many thanks to my trombone teacher Frank Szathmáry-Filipitsch, who helped me overcome individual challenges regarding articulation on the trombone (no doubt informing my interest in the subject) by showing me how to effectively teach myself and others.

Around this time, I also began parallel studies in secondary school teaching; my choice of English as the required second subject (to complement Music) was more or less a matter of convenience at this stage – little did I know where this journey would take me! Bill Barry taught English phonetics and phonology to English language majors at Saarland University at the time, and it was to him that I first mentioned my unusual hypothesis. He immediately seemed to realize it was a topic that would require substantial enquiry (at the time I was considering doing it as a Staatsexamen-thesis – a 6-month project!), and he later gave me some feedback on the research proposal that I submitted to the University of Canterbury with my scholarship application. In the meantime, however, my interest in the topic had faded in favor of other interests in Language and Music that I developed while working as a research assistant for Neal Norrick. My Staatsexamen thesis was concerned with the interaction of musicians in rehearsals, reflecting the discourse-based approach prevalent in the department; the brass playing topic only resurfaced when I decided to try pursuing a Ph.D. in Language and Music rather than moving on to teacher training. I emailed University of Canterbury Linguistics staff about applying for a Doctoral Scholarship, and even though my initial topic idea did not prove feasible, the forgotten research question quickly emerged as the perfect fit when asked whether I might be able to do a project involving ultrasound imaging of the tongue.

I would like to express my immense gratitude to my supervisors Jen Hay and Donald Derrick, who not only guided me through the Ph.D. process but let me convince them that the proposed topic was worthwhile of study. I am sure they uttered a huge sigh of relief when they were first able to observe ultrasound video of the pilot participant playing the trombone; of all my fantastic participants, this person deserves individual praise for 'suffering' through the whole procedure twice, the second time around showing that double-tonguing on the trombone is possible even with three articulography sensors glued to the tongue!

I would like to acknowledge all staff and fellow students in the Department of Linguistics and at the New Zealand Institute of Language Brain and Behaviour at the University of Canterbury for all their generous support, and their helpful comments provided at various 'Socio' meetings (weekly lab meetings). Kevin Watson was my first contact when considering to apply for a UC Doctoral Scholarship and instrumental in re-discovering my unusual research idea that turned out to be such a great fit. Heidi Quinn and Susan Foster-Cohen always showed great interest in my musical pursuits outside university, attending a number of concerts I participated in, as well as never complaining about the trombone playing coming out of my office at night; a big thank you also goes out to a number of other people who, at various times, put up with my playing, especially the fellow late-night workers in the office on the other side of the wall! A number of people were instrumental in helping me acquire the skills necessary to use the various computer programs I required for my analysis. James Gruber and Jacqui Nokes helped me get started in Praat, while Romain Fiasson and Tom de Rybel provided help with MATLAB and other programming language-based applications. The latter two, along with external support from Mark Tiede (Haskins Laboratories), were also instrumental in helping me overcome various challenges posed by trying to record ultrasound video at maximum frame rate; this included mastering the Windows command line and FFmpeg, and issues with video formats and their compatibility with all the different programs I needed to use. Vica Papp was of great help with advanced Praat questions (particularly on side projects on NZE), and Pat LaShell and Jen Hay's enthusiasm provided early motivation to dig into statistics that proved vital when getting into more advanced stats and data manipulation in R. I always enjoyed the conversations I had with Clay Beckner about various parallels of language and music, and similar discussions with Ryan Podlubny and Andy Gibson once they joined the department and started their own PhD projects linking the two fields. Daniel Bürkle, Ahmad Haider, Mohammed Dagamseh and Moonsun Choi were wonderful office mates, and fun times were had with a cohort of ‘older’ and similar stage Ph.D. students and post-docs (Petya Racz, Mari Sanchez, Jiao Dan, Annalise Fletcher, Ksenia Gnevsheva, Darcy Rose, Keyi Jason Sun and Xuan Wang), as well as some more recent arrivals (Arshad Ali, Khalid Aljawazneh, Daiki Hashimoto, Jacq Jones and Mineko Shirakawa).

My thanks also go to the UC library interloans staff who procured a total of 120 items for me, including 17 theses; NZILBB and our wonderful staff Emma Parnell and Scott Lloyd for letting (and helping) me use all kinds of equipment, most importantly the ultrasound machine; Robert Fromont for LaBB-CAT support; Warwick Music UK for sending me a free ‘pBone’; and of course the University of Canterbury for awarding me a Doctoral Scholarship that ensured my survival throughout this thesis, price money for winning the College of Arts Thesis in Three competition, and several small research and travel grants awarded by the School of Language, Social, and Political Sciences. Finally, Morgana Mountford-Davies and Ryan Podlubny deserve the author’s and readers’ gratitude for a great job of proof-reading and cutting down on my often all too Germanic sentences.

Jim Scobbie and Murray Schellenberg were great examiners and I am indebted to them for agreeing to examine a thesis on an unusual topic as well as providing very helpful comments that made this thesis a better document.

Outside academia, I am very grateful to my parents who have supported me throughout all my studies, and my wonderful friends who made sure I kept myself busy during my Ph.D. by playing music and getting out into the fantastic NZ bush.

Vita

2012	State Exam, “Teaching for secondary schools” (1. Staatsexamen) with the subjects Music and English, University of the Saarland and University of Music Saar, Saarbrücken, Germany
2012	Diplom, “Jazz – complementary studies” Trombone (Diplom-Ergänzungsstudiengang Jazz), University of Music Saar
2010	Diplom, “Orchestral Music” (performance) Bass Trombone, University of Music Saar

Abstract

This thesis attempts to answer the question whether there is an influence of native language on the playing of brass instruments. Two hypotheses are formulated based on an extensive literature review of brass instrument acoustics, previous empirical research, and findings from motor control and speech production. Hypothesis 1 posits that brass players can perceive differences in the playing of performers from different language backgrounds, and is cautiously answered in the affirmative by findings from an online questionnaire that shows that players believe they can perceive such differences. Hypothesis 2a) predicts that the tongue position assumed during sustained note production on brass instruments is based on motor memory from a player's native language vowel production, with an extension b) predicting that functionally independent sections of the tongue would pattern individually and be affected differently by language influence, offering support for modular accounts of motor control. To address this hypothesis, an ultrasound study of ten Tongan and nine New Zealand English-speaking trombone players is carried out, recording participants while reading wordlists and during trombone playing. Results show clear differences between the average tongue positions employed by performers from each group. Except in a few individual cases, there is no match between the overall tongue shape for vowels and sustained note production on the language group level; different patterns apply for the back and front of the tongue. This finding supports the extension of hypothesis 2 and provides evidence for modular theories of motor control and their application to the vocal tract musculature. Various constraints related to airflow, acoustics, and articulatory efficiency are discussed; it is suggested that language influence, while clearly visible in the results, is secondary to these constraints. Confounds of the study include the difficult nature of ultrasound probe stabilization during trombone playing, the challenge of comparing articulatory movements during two very different activities, and differences in trombone playing proficiency across the two observed language groups. In addition to providing support for modular theories of (speech) motor control, the thesis makes an important contribution to ultrasound methodology by proposing a principled way for normalizing across participants and different vocal tract activities.

1 Introduction

The use of speech syllables in teaching brass instruments can be traced back to as early as 1584 (Dalla Casa, 1584/1970) and brass teachers have long used different consonants (/t/ versus /d/ for hard versus soft attacks) and vowel colors (/ɑ/ versus /i/ for low versus high range notes) to illustrate what students should do with their tongue to produce desirable sounds on brass instruments. Furthermore, anecdotal accounts of language influence on brass playing have been exchanged within the brass community, for example, speculation about why players of some nationalities are 'better' than others at certain facets of brass playing or why learners may have specific problems related to their language background. An old, but classic, example of the former is Fitzgerald's (1946) report of the great cornet soloist Herbert Clarke's thoughts about 'Latin' brass players: "their language may help them to be more decisive, besides guiding them with greater certainty as to the attack for the different varieties of tongueing" (as cited in Fitzgerald, 1946, p. 5). Fitzgerald tried to add credibility to Clarke's speculation by commenting that "[t]his opinion is well founded since the Latin language and those closely related to it employ a much greater variety of vowel sounds than the average American uses in his speech and requires both extreme flexibility and velocity in lingual movement, particularly in the use of the tip of the tongue" (pp. 5-6). Note, however, that linguists agree that Latin languages (understood here to be Spanish and other Romance languages) have *fewer* vowel phonemes than American English. Possibly even more perplexing to the informed reader might be the following quote by Jean Devémy, a famous French horn player and teacher from the Paris Conservatoire, printed originally in a French periodical:

> Everyone knows that the main point of horn technique consists of perfecting the tone quality. In comparing French horns with German horns, everyone is aware that there is a striking difference between them from the point of view of tone quality. This difference, contrary to what is generally believed, does not in any way originate from the bore or from any other technical details of workmanship. A horn manufactured in Erfurt and a horn manufactured in Paris are not notably different. It is only the position of the lips, the structure of the throat of the

> performer, *due to the language of his country*, which makes the difference in sound (as cited in Barboteu, 1975/2000, p. 35; emphasis in the original)[1].

Focusing on the problems of certain populations of players, one can find a small number of more recent and linguistically informed accounts of possible language influence, such as Joseph Bowman sharing his experience of teaching trumpet to Thai students in the *International Trumpet Guild Journal* (Bowman, 2011): "Looking at the Thai language specifically, the wonderful tonal language contains very few hard consonants. A hard "taa", "kaa," or "gaa" sound doesn't exist, so introducing those takes time and persistence" (p. 90). As any brass player would know, these are the kinds of syllables that most teachers and methods advocate using when articulating on a brass instrument.

The assumptions underlying such assessments and the use of certain syllables, however, have rarely been tested. It is well-known that the different native languages spoken by brass players would have different plosive articulations and vowel inventories, but I do not know of any empirically-grounded research predating my study on whether a player's native language might affect the vocal tract configuration used, and thus, the sound produced, during brass playing (there exist two Doctor of Musical Arts dissertations addressing this question, albeit with methodological shortcomings; these will be discussed in section 2.7).

This thesis attempts to answer the aforementioned question with empirical data of tongue positioning collected using ultrasound imaging, and is structured in the following way: chapter 2 provides background information on brass instrument acoustics and the influence of vocal tract resonances on brass instrument sound, before moving into an overview of suggestions provided in historical and contemporary method books regarding tongue position during brass instrument performance. This is followed by a review of previous empirical research on brass playing to provide the reader with a thorough overview of the underlying physics, pedagogical concepts and vocal tract movements affecting contemporary performance on brass instruments. The chapter concludes with two short sections on the perception of brass instrument sound

[1] Barboteu, himself a renowned French horn teacher from France, shows that he agrees with Devémy's judgment by putting the following sentence directly after the quotation: "It is true that the spoken language of a country can give individuality to a hornist's playing" (Barboteu, 1975/2000, p. 35); he does not add any clarification, however.

by (expert) listeners, and a summary of the two existing dissertations on language influence on brass playing.

The literature review continues in chapter 3 on the physiology and motor control of the upper vocal tract during speech production and brass playing. This chapter examines the contributions of selected articulators to both activities, and discusses various models of motor control that could underlie the observed vocal tract movements. A specific focus is placed on modular accounts of motor control and the predictions such an approach makes regarding cross-system interactions between the two activities.

Chapter 4 outlines possible areas of language influence on brass playing, and presents the two hypotheses informing the research carried out in this thesis: (1) Brass players can perceive (consciously or subconsciously) the acoustic consequences of playing differences between players with different native languages. And (2), tongue positions assumed during sustained note production on brass instruments are based on motor memory, (a) where such motor memory is based on speech articulation, specifically the tongue shape for vowels. Furthermore, it is hypothesized that functionally independent sections of the tongue will be individually affected by motor memory from a player's native language, in agreement with a modular theory of motor control (hypothesis 2b).

Results from an online questionnaire completed by 135 brass players world-wide are presented in chapter 5. Participants' responses suggest that brass players at least believe that they can perceive differences among the performances of players with different native languages, providing limited support for hypothesis (1). While future research is clearly needed to properly address this question, the questionnaire data provide an important link between the potential of vocal tract influence on brass instrument sound (upon which language influence is dependent) and its audible consequences.

The remaining chapters document the ultrasound study conducted to address the main research question of this thesis, whether native language influences sound production on brass instruments. A brief account of ultrasound imaging of the tongue as the chosen methodology is presented in chapter 6.

The following chapter describes the data collection from two groups of trombone players whose native languages were New Zealand English (NZE) and Tongan. Participants were asked to read wordlists in their native language before playing selected musical passages, which allowed me to record their midsagittal tongue positions during both activities.

In chapter 8, various steps of the analysis are described, which involved finding a solution to normalizing ultrasound data across individuals.

Pooled and individual results of this process are presented in chapter 9, along with a measure allowing the quantification of differences between midsagittal average curves. While the average tongue contours during sustained note production on the trombone are clearly different for the two language groups, they do not pattern with any overall vowel tongue shape for either language. Separate comparisons for the back and front of the tongue, however, provide support for hypothesis (2b): the position of the back of tongue during note production patterns with the back vowels for each language, while a vowel-unrelated constraint seems to lead to a difference at the front of the tongue. At the end of this chapter, I also present limited place of articulation data for coronal consonants produced during speech, and coronal attacks produced during trombone playing, for one player from each language group.

Chapter 10 offers a discussion of the findings of this study, listing various non-linguistic constraints that affect tongue position during brass playing and interact with the influence of a player's native language. Language influence, though secondary to other constraints, seems to have a non-negligible impact on tongue position, and affects the back and front of the tongue differently. Possible explanations for this finding are provided and the chapter closes with implications for brass playing and teaching, and suggestions for further research.

Chapter 11 provides a conclusion to round out the thesis.

2 The Acoustics and pedagogy of brass instruments and previous research

Although the principles of brass instrument acoustics are not yet fully understood, recent research suggests that there is some influence of the vocal tract on brass instrument sound. Vocal tract influence, of course, forms an important prerequisite for language influence on brass instrument sound, and the acoustics research reviewed in this chapter suggests that there are a number of areas of brass playing that could be affected by a player's native language (L1) via its articulatory movements and default settings. A comprehensive overview of previous empirical research on brass playing suggests that language influence likely involves the tongue but possibly also the larynx and alterations of the shape of the pharynx. Unique articulatory features of a particular language would be acquired early on in life as part of one's L1, and their influence is unlikely to be noticed consciously by brass players. This may well be the reason why the possible connections between speech and brass playing have received little empirical investigation so far, despite a long tradition of using speech syllables in brass teaching. Furthermore, the use of such syllables in the pedagogical literature will be outlined from the earliest book on recorder playing (1535) to the present, showing a clear focus on 'cardinal' or orthographic vowels, and problems that may arise when such syllables are used by players who do not share the L1 of the author.

2.1 Brass instrument acoustics: A simplified model

In a simplified model, tone production on brass instruments can be regarded as happening via an outward-striking lip-reed mechanism (often referred to as the player's 'embouchure') that excites the air column within the instrument, producing a spectrum of standing waves which are controlled by the natural frequencies of the air column and which are emitted from the bell at varying volumes (cf. Campbell & Greated, 1994; Benade, 1976). Figure 2.1a on the next page provides a representation of such a simplified model, along with a so-called 'water trumpet' in figure 2.1b, which uses an "open channel filled with water" to depict (sound) wave oscillations within lip-reed instruments (Benade, 1976, p. 392). In figure 2.1b, a float-operated valve ("player's lips" in figure 2.1a) regulates the periodical water flow into the trough, opening "progressively as the water level rises at the 'mouthpiece' end of the trough" and decreasing flow "when the water level is low" (Benade, 1976, p. 392); this mechanism serves to maintain oscillation (a standing wave) in the water channel.

Other features of brass instruments such as the bore taper and bell flare are represented in the water trumpet by increasing water depth (attenuating wave amplitude), and the vertical line restricting leakage at the right end of the channel.

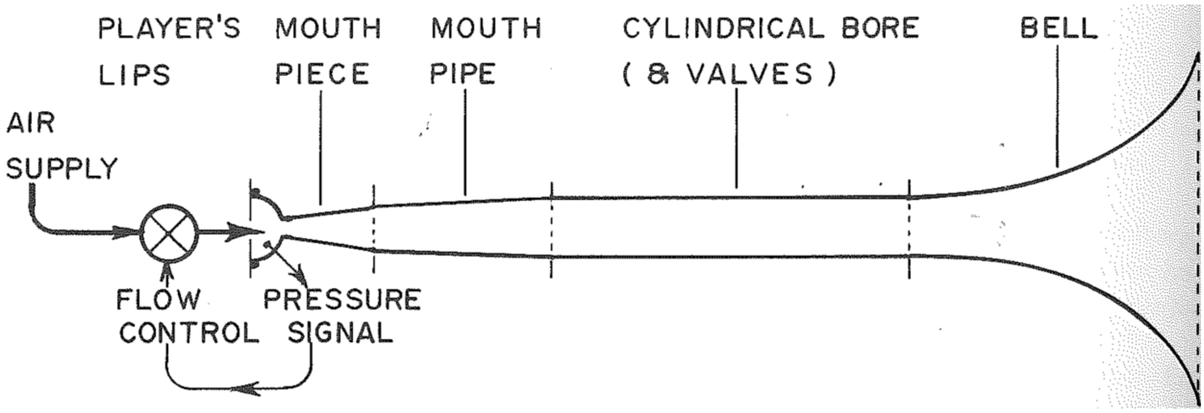


Figure 2.1a: A simplified model of a lip-reed instrument.

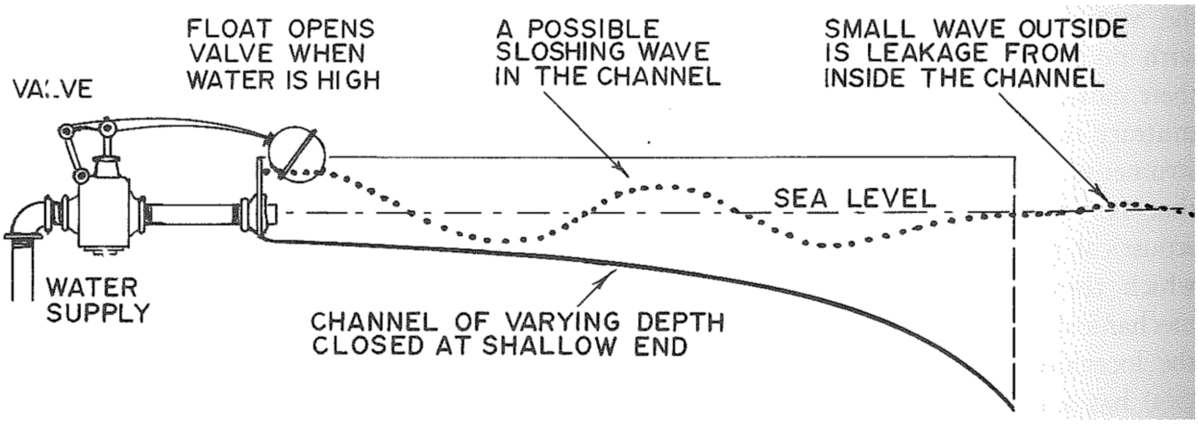


Figure 2.1b: ‘Water Trumpet’ using a water-filled channel to depict wave oscillations within lip-reed instruments. From Benade, 1976, p. 392.

Another option for simplification is to regard brass instruments in analogy to the human voice; here, one can think of the embouchure as the brass player’s larynx and the instrument as the vocal tract, which serves merely as an amplifier. The number of close pitches that can be produced, however, is much smaller than for the human voice due to the much greater length of instrument as compared with the vocal tract. The mechanical shape of a brass instrument rather “has a direct influence on the shapes of the puffs of air which enter its mouthpiece” and thus does “not only […] transmit sound components selectively from the flow source to the room, it also plays a large role in determining the nature of the incoming flow pattern itself” (Benade, 1976, p. 391). Consequently, there exists no mechanism comparable to the human tongue that would enable brass players to produce such as wide range of timbres as we do in speech and singing[2].

In more scientific terms, the embouchure (or lip reed) can be described as

> a form of valve which allows the high pressure air from the player's mouth to enter the instrument in a series of pulses which inject energy to initiate and maintain the oscillations within the tube. The oscillations of the reed are strongly influenced by the oscillations within the air column, i.e. there is a strong coupling between them (Campbell & Greated, 1994, pp. 259-260).

This explains why the physical dimensions of a brass instrument very much determine the notes that are playable for a given length of tubing. Figure 2.2 on the following page shows waveforms of the pressure variations in a trombone mouthpiece (left column) for four different notes increasing in pitch from top to bottom, and the resulting harmonic spectra of the resonated sound (right column). Note that these complex tones are made up of oscillations at various frequencies (determined by the natural overtone series), the number of which diminishes with increasing pitch.

[2] Hézard et al. (2014) carried out a systematic study on *Synchronous multimodal measurements of the lips and glottis*, describing the differences between both systems in more scientific terms: "… the coupling between the lips an [sic] the air column can be very strong in brass instruments, primarily due to the geometry of the downstream resonator …, whereas it is much more moderate in voice production. In other words, the strong impedance resonances of a brass instrument bore only allow in theory a finite number of frequencies to be produced for a given configuration of the air column" (p. 1173).

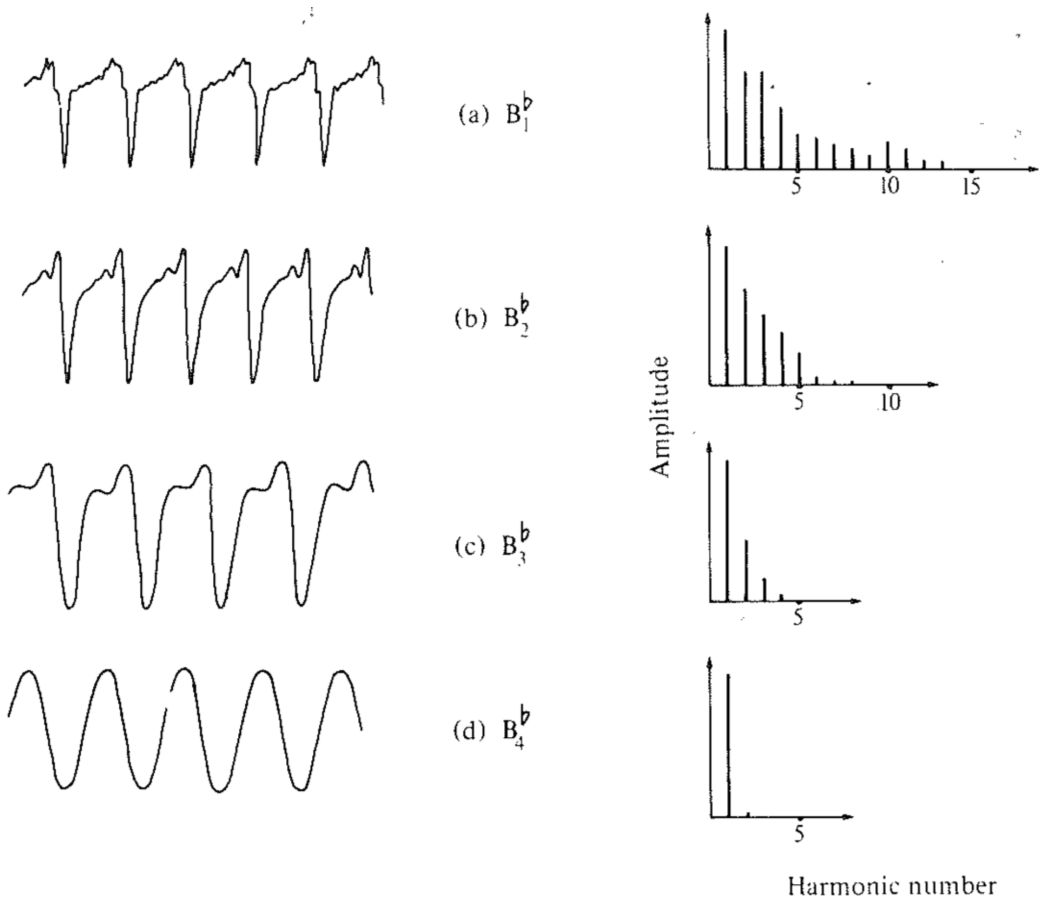


Figure 2.2: Waveforms (left) and harmonic spectra (right) of the pressure variations in a trombone mouthpiece during the playing of four notes. Reproduced with permission from Campbell & Greated, 1994, p. 321.

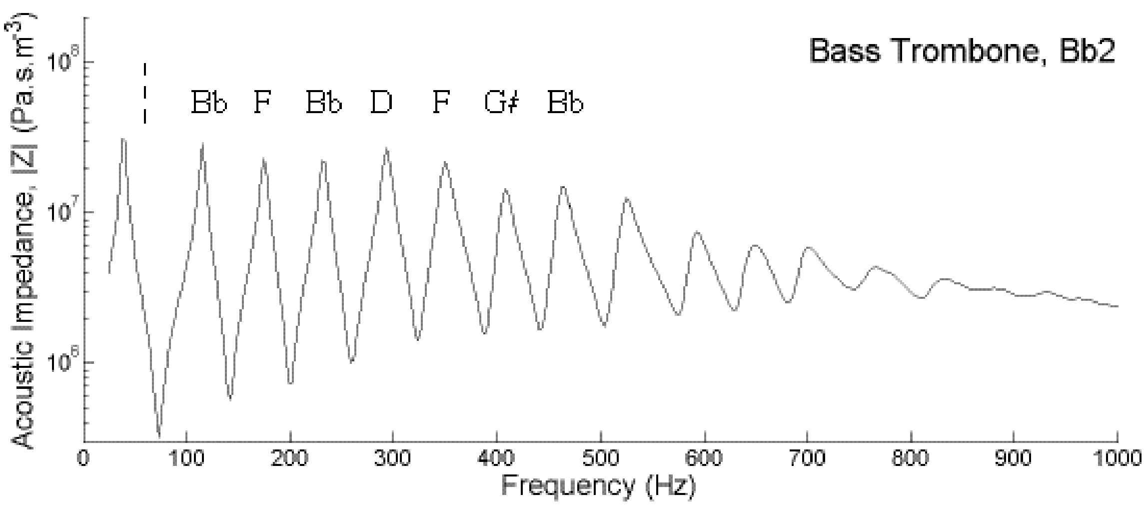


Figure 2.3: Impedance spectrum for a Bb bass trombone (slide in first position, valve not depressed). Reproduced with permission by the creator of the image (Jer-Ming Chen) from http://newt.phys.unsw.edu.au/jw/brass/brassZ/bass_trombone_Bb2.gif, last accessed 21 September 2016.

Figure 2.3 on the previous page provides an example of the impedance spectrum for a Bb bass trombone with the slide in first position and the valve/s not depressed. Acoustic impedance for tubes is commonly defined as the pressure amplitude for a certain wavelength of oscillation divided by the attenuation of the tube, which is related to its diameter; of course, it is also related to the tube's length, as this constrains which oscillations can form standing waves within the tube. An impedance spectrum indicates the level of damping affecting standing waves at various frequencies, where peaks correspond to minimal attenuation while valleys indicate significant damping. The peaks on the spectrum displayed in figure 2.3 are the overtones or higher partials of the fundamental (Bb2) for the given instrument length (the extension of the slide or engagement of valves can alter the fundamental frequency) and indicate the pitches that can be played by adjusting the vibrating frequency of the lips. Note that even though the distances between peaks appear to be equidistant, perceptually they correspond to the narrowing intervals of the harmonic series with increasing pitch (octave, fifth, fourth, major third, etc.) due to the roughly logarithmic perception of pitch (see figure 2.4 on the following page). Professional brass players are highly skilled at precisely manipulating the vibrating frequency of their lips to select the different partials available for a certain tube length; however, departing significantly from the optimal pitch pre-specified by the length of the instrument will usually cause the lips to re-adjust their vibrating frequency by shifting to an alternative impedance maximum, resulting in a 'split' note. An added difficulty for brass players is "the fact that it takes a long time for acoustical 'messages' to travel from mouthpiece to bell and back, informing the lips of the collaborative job they must do with the air column" (Benade, 1976, p. 425; drawing on Benade, 1969) which complicates achieving a clean attack and changing notes. This can additionally be complicated

> by a small change in cross section, a sharp bend, or an ill-chosen change in the taper […]. Such discontinuities return a premature echo of significant size to the mouthpiece, an echo that is not even a replica of the original disturbance. Such ill-timed, ill-shaped return echoes can upset the best-trained of lips, and, having spoiled the steadiness of their initial vibration, will ruin the attack (Benade, 1976, p. 425).

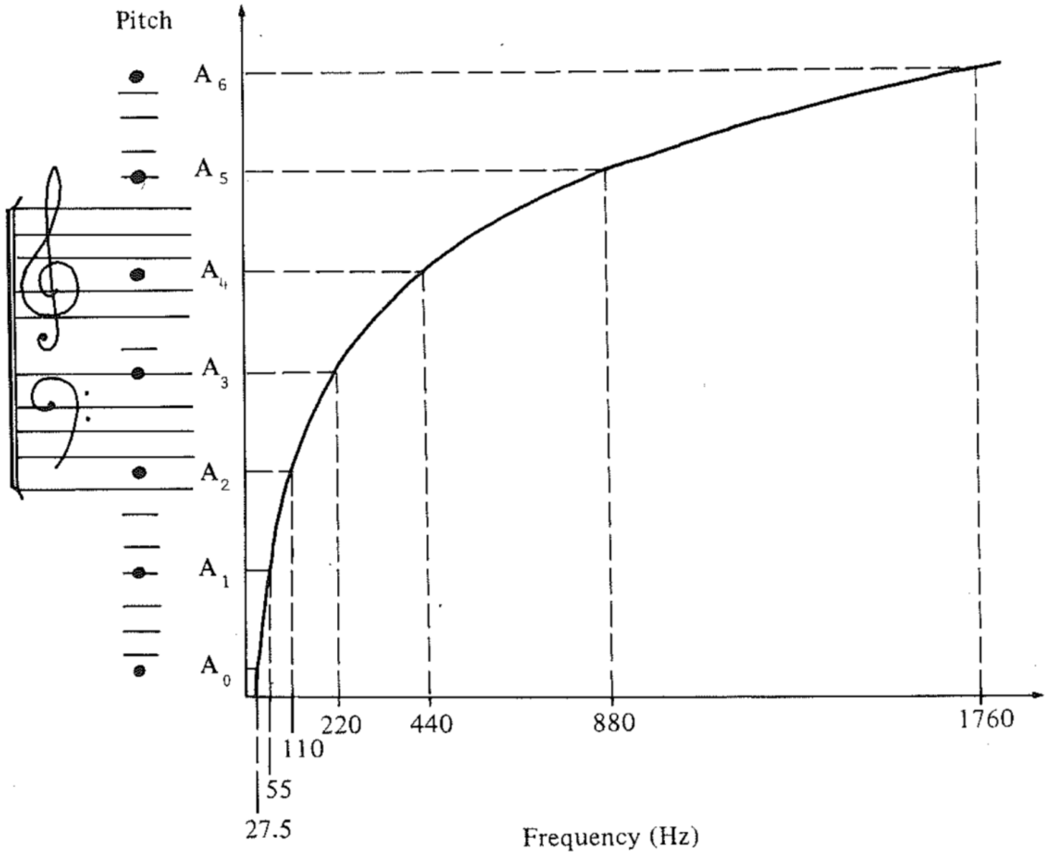


Figure 2.4: The relationship between pitch and frequency. Reproduced with permission from Campbell & Greated, 1994, p. 75.

2.2 Further considerations

The impedance spectrum for a specific bass trombone plotted in figure 2.3 is the result of interactions between the transfer functions of different sections of the bore making up the instrument, as well as other factors that are harder to estimate, such as acoustic losses associated with different materials used in instrument manufacturing. Impedance curves for different brass instruments generally look quite similar, albeit varying predictably along certain dimensions relating to the acoustic characteristics of their various constituent parts. Moreover, the spectrum illustrates why it becomes increasingly more difficult to produce notes in the high range of a brass instrument: not only are the available peaks closer together (in terms of intervals), increasing the chance of hitting an unintended note, but the fact that natural tones are always made up of multiple partials (fundamental plus overtones at varying multiples of the fundamental) also means that there are fewer or no peaks to support the oscillations of the higher partials of the composite tone (cf. figure 2.2).

The reason why the peaks in figure 2.3 become increasingly smaller, eventually dying out completely, is the so-called cut-off frequency applying to any brass instrument with a flared bell section, illustrated in figure 2.5 below. This frequency lies at around 700 Hz for the trombone (Campbell & Greated, 1994, pp. 346-347). High frequency waves exceeding the "forbidden zone" (shaded area in Figure 2.5, cf. Campbell & Greated, 1994, p. 346) do not get reflected back into the instrument and thus are almost completely transmitted by the bell; conversely, the lack of reflection means that no stable standing waves can be initiated at those frequencies. For lower pitches, the barrier ("forbidden zone") becomes increasingly "thicker and higher" and thus a "smaller […] fraction of the internal sound" manages to "tunnel[s] […] through to the outside world" (Campbell & Greated, 1994, p. 347).

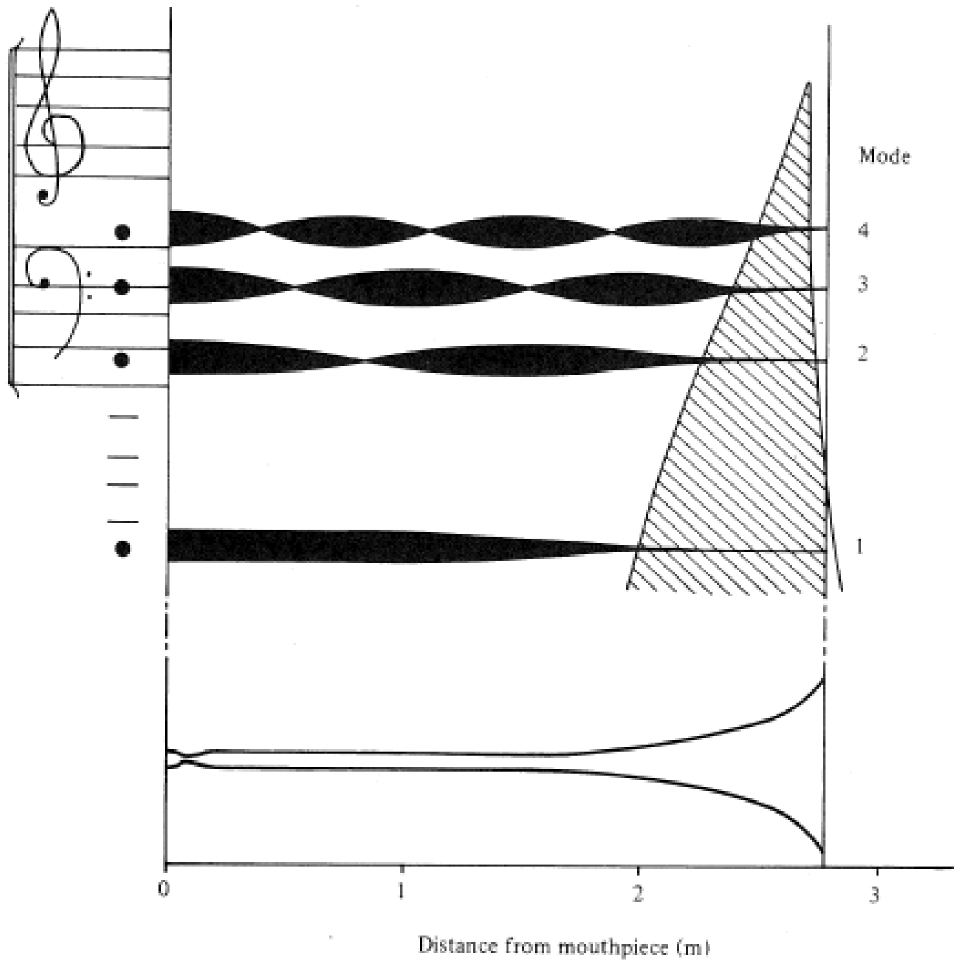


Figure 2.5: "Standing wave patterns for the lowest four resonant modes of a Bb tenor trombone. The shaded area represents the 'forbidden zone,' which reflects most of the low frequency sound energy arriving at the bell." Reproduced with permission from Campbell & Greated, 1994, p. 346.

While the simplified model described above can explain basic features of sound generation on brass instruments, more recent research has shown that it is insufficient to account for more advanced playing techniques. Evidence comes from studies observing the brass player's lips in motion and attempts at building and/or modeling artificial lips (Adachi & Sato, 1996; Bromage, Richards, & Campbell, 2003; Cullen, Gilbert, & Campbell, 2000; Ludwigsen, 2003). Aside from the simplest model where the lips are blown open when mouth pressure increases (the aforementioned outward striking mechanism), there also exist the possibilities of the lips being "blown laterally (in the higher range), or by some combination of these" (Chen, Smith, & Wolfe, 2012, p. 722; drawing on Yoshikawa, 1995; Adachi & Sato, 1996; Yoshikawa & Muto, 2003). This allows the player "to produce notes above, below and at the resonant frequencies of the instrument" (Bromage, Richards, & Campbell, 2003, p. 197). Although I will not review this line of research in more detail, it is closely tied to the issue of vocal tract influence on brass instrument sound discussed in the following section.

All of the acoustical parameters discussed so far behave more or less linearly at low and medium dynamics, with the strength of the overtones increasing faster than that of the fundamental as the instrument is blown harder/louder (Benade, 1978, p. 51; drawing on Worman, 1971). Above a certain threshold, however, further increasing the sound pressure leads to variable temperature rises within the instrument according to the existence of pressure maxima and minima; this in turn leads to inconsistencies in the speed of sound throughout the bore and a distinct change in timbre referred to as 'brassy sound' (cf. Chick et al., 2012; Norman et al., 2010). Myers et al. (2012) came up with a measurement to determine the "brassiness potential" of different types of brass instruments and found that the amount of cylindrical tubing affects this potential, making a small bore tenor trombone the instrument with the highest such potential. In extreme cases, nonlinear sound propagation might even lead to the generation of shock waves in the bore (Hirschberg et al., 1996). Temperature changes and the concentration of $C0^2$ within the tube also affect playing characteristics at lower dynamics, though to a smaller extent, as documented by Boutin et al. (2013). These authors performed measurements of the relevant parameters following varying durations of playing, 'flushing' the air inside their test instrument between trials (Boutin et al., 2013, p. 3).

2.3 Vocal tract influence on brass instrument sound

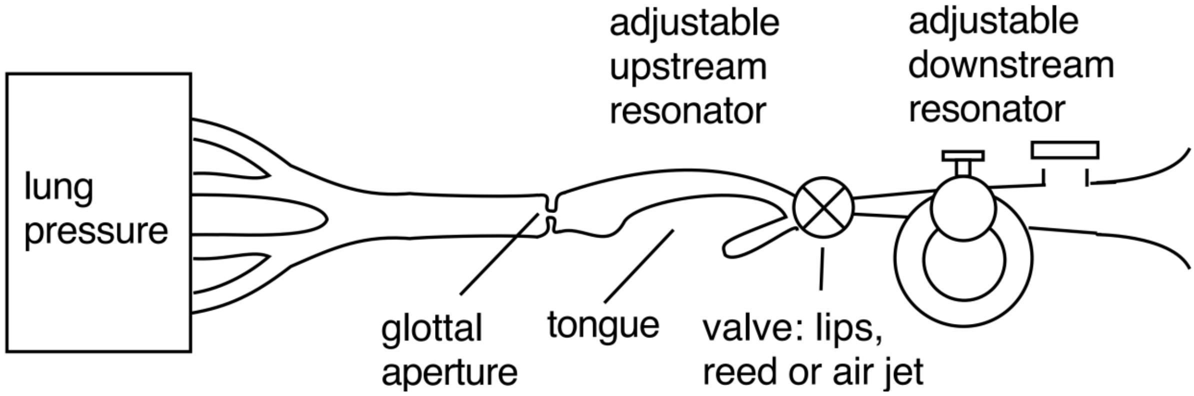


Figure 2.6: “A simplified schematic (not to scale) showing most of the elements controlled by the player, beginning with the pressure of the air in the lungs. The adjustable glottal aperture admits air into an upstream duct of adjustable geometry, including the possible constriction or occlusion by the tongue. Sustained oscillation depends on adjustments to the valve – the air jet, reed or lips – whose parameters are carefully controlled. The geometry of the instrument bore is adjusted by valves, keys or a slide (the last not shown), and the bell, if present, may be modified by the hand or mutes.” Reproduced with permission from Wolfe, Fletcher, & Smith, 2015, p. 1; first printed in Hanna, Smith, & Wolfe, 2012.

The potential of language influence on brass playing would be very limited if there was no interaction between the resonances of the vocal tract and the oscillations within the instrument. Figure 2.6 above illustrates various parameters under the control of a wind instrument player, including the shape of the vocal tract (represented by the tongue), and the flow into (glottal aperture) and out of it (via the valve, or lip reed). While Clinch, Troup, and Harris (1982) were among the first to investigate the phenomenon scientifically, Benade (1985 & 1986) proposed one of the earliest models for “[i]nteractions between the player’s windway and the air column of a musical instrument” (title of his 1986 paper; see also less formal descriptions by Stauffer, 1954; Coltman, 1973). Of course, the musicians themselves had long insisted on such a mechanism (Benade, 1985 & 1986 duly acknowledges this), with previous investigations by acousticians disagreeing on whether “vocal tract resonance frequencies must match the frequency of the required notes” (Clinch, Troup & Harris, 1982, abstract) or that “the effect of the player's vocal tract on the instrument's tone quality should be negligible” (Backus, 1983, abstract). In possibly the earliest textbook to mention the possibility of vocal tract influence on wind instrument sound, Campbell & Greated (1994) describe the phenomenon as follows:

> As long as the lungs, throat and mouth are treated purely as a means of supplying air at constant pressure behind the lips, it is difficult to see why their shape should be relevant. However, we can look on the player's windway as effectively a second 'brass instrument', with the air flowing in the 'wrong' direction (that is, towards the lips rather than away from them) ... The tubes and cavities of the windway will also have an acoustic impedance, and therefore the fluctuations in the air flow introduced by the lip vibration will cause a fluctuating pressure difference between the lungs and the mouth (pp. 324-325).

Recently, documentation of the necessity of tuning one's vocal tract to sound the notes within the altissimo register of saxophones (Chen, Smith, & Wolfe, 2008 & 2012; Scavone, Lefebvre, & da Silva, 2008) and to achieve the clarinet glissando in the first bar of *Rhapsody in Blue* (Chen, Smith, & Wolfe, 2008), as well as investigations into the playing technique of the didgeridoo (Fletcher, 1983; Tarnopolsky et al., 2005), have shown that vocal tract resonances are quite important for the production of wind instrument sound (see also Johnston, Clinch & Troup, 1986; Fritz et al., 2003; Fritz & Wolfe, 2005; Coelho & Iazzetta, 2011). Even if vocal tract resonances do not determine the fundamental pitch of the produced note, they might influence its timbre (Scavone, Lefebvre, & da Silva, 2008; Wolfe, Garnier, & Smith, 2009; Li et al., 2015; see also section 2.3.1.1 below). A schematic representation of how one's vocal tract resonances might influence wind instrument sound is given in figure 2.7 on the following page for brass (lip-valve) instruments (b) and reed instruments (such as clarinet & bassoon) (c).

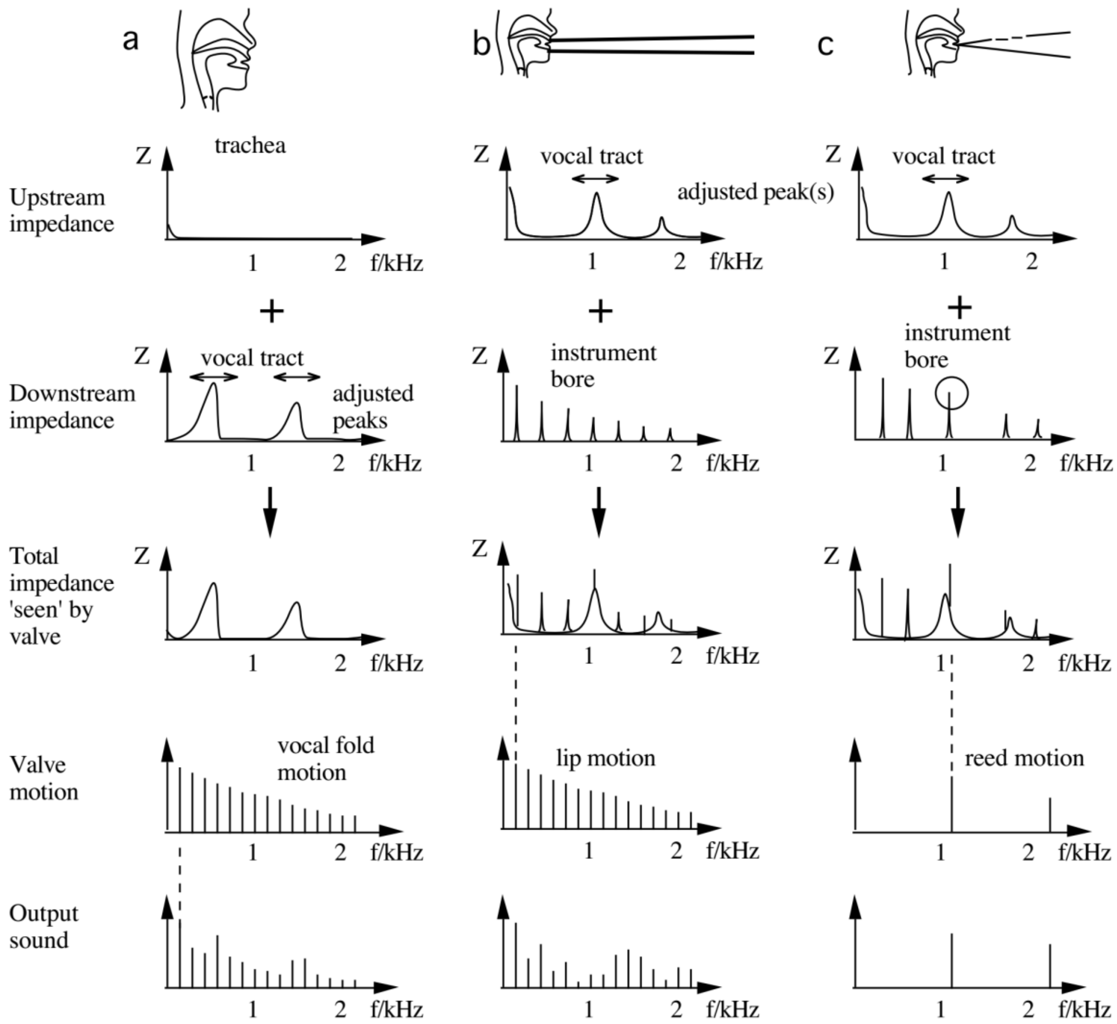


Figure 2.7: “Schematic figures show idealisations of the duct-valve interactions in the voice (a), a lip valve instrument (b) and a reed instrument (c). The upstream and downstream impedances add to give the total impedance Z. (a) illustrates the Source-Filter theory: the vocal fold motion is assumed independent of Z, but tract resonances transmit certain harmonics more efficiently into the sound field at the open mouth. The vocal tract impedance spectrum has a different form when seen from the glottis (a) and from the lips (b) and (c). In (b) and (c) the mouth is closed so it will be difficult to change low frequency impedance peaks significantly by altering the mouth geometry. An impedance peak may be changed over a range near 1 kHz, which alters the total impedance (instrument + tract) over the same range. The timbre effect, as used by didjeridu players, is shown in (b) where impedance peaks inhibit output sound. For a high pitched instrument, such as a saxophone played in the altissimo range (c), a resonance may help determine the playing range by selecting which of the sharp peaks in the impedance spectrum of the instrument bore will determine the playing regime.” Reproduced with permission from Wolfe, Garnier, & Smith, 2009, p. 4.

Observing the playing behavior of “an artificial trombone playing system” employing “highly simplified models of lip, vocal tract and glottis,” Wolfe, Chen, and Smith (2010)[3] report that the system “played sharper” when used with a “high tongue configuration […] than the low tongue model” (p. 7; revisiting findings from Wolfe et al., 2003). This difference was observed without changing instrument length, meaning the artificial ‘lip reed player’ (a variably shaped cavity combined with a cantilever spring) “operated on the same impedance peak of the bore. As the slide was extended, there was also a range over which the high tongue model played on a higher resonance” (Wolfe, Chen, & Smith 2010, p. 7), producing a differently pitched note (see empty circles in figure 2.8 below; cf. similar findings for the clarinet in Fritz et al., 2005). While the authors report similar observations by skilled players (lowering the tongue during sustained note production made the pitch go flat or drop to the next lowest register), they stress the advantage of observing an artificial player system that does not automatically make adjustments to counteract the pitch changes caused by an altered tongue position (Wolfe et al., 2003, p. 310). They conclude that “raising the tongue, or the tongue tip, increases the height of peaks in the vocal tract impedance, and so more effectively couples it to the instrument resonances and to the reed or lips. This gives wind players a method of fine pitch adjustment, by variably coupling a largely imaginary impedance” (Wolfe et al., 2003, p. 310).

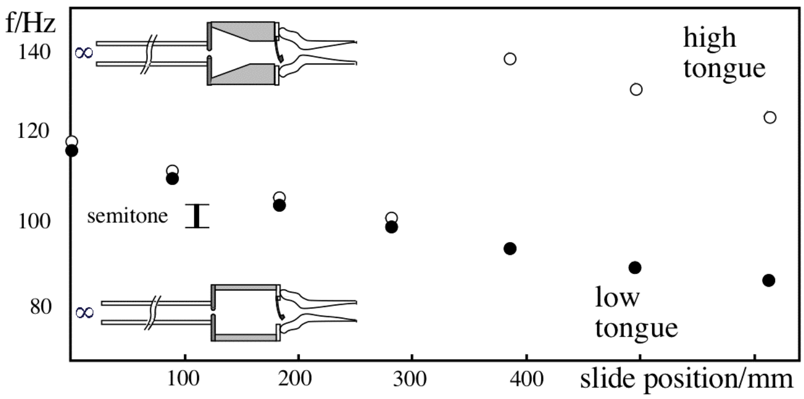


Figure 2.8: “The playing frequency of an artificial trombone playing system as the slide is extended from the closed position (0 mm). The open [and] filled circles refer to geometrically simplified ‘vocal tracts’ represented by the sketches for ‘high tongue’ (top) and ‘low tongue’ (bottom).” Reproduced with permission from Wolfe, Chen, & Smith, 2010, p. 7 (first printed in Wolfe et al., 2003).

[3] This paper discusses vocal tract influences on various wind instruments.

2.3.1 *In vivo* measurement of vocal tract influence on brass instrument sound

A small number of studies have tried to investigate the influence of vocal tract tuning on brass instrument sound in human subjects, using two different modes of investigation. One way to do this is to determine simultaneously the pressures inside the mouth and the mouthpiece, based on the assumption that the flow into the bore at each frequency has to be the opposite of that into the mouth (cf. Elliott & Bowsher, 1982; Benade, 1985). Comparing the measurements at both locations gives “the ratio of the acoustic impedance of the two ducts” (Wolfe et al., 2015, p. 10; drawing on Wilson, 1996; cf. Li et al., 2015). The advantages of this method are its “speed and relative simplicity so that the time variation in this ratio in performance can be measured” (Wolfe, Fletcher, & Smith, 2015, p. 11); however, the drawback is that it only samples frequencies at the harmonics of the note played. This method has been employed by a team of researchers at McGill University in Canada, and I will discuss their studies on the role of vocal tract influence on trombone playing below. The alternative is “to measure the impedance spectrum in the vocal tract by injecting a known broadband acoustic current into the mouth” (Wolfe et al. 2015, p. 11). In this case, notes have to be sustained for roughly a second, but this makes it possible to directly determine the resonances of the vocal tract.

2.3.1.1 Direct measurement of vocal tract influence on brass instrument sound

A team of researchers at the University of New South Wales (UNSW) in Australia has applied the setup described above to measure vocal tract influence on saxophone, clarinet, didgeridoo (see references listed above), trumpet (Chen, Smith, & Wolfe 2012), and trombone performance (Boutin et al., 2015). Their study on trumpet playing (Chen, Smith, & Wolfe, 2012) found that the measured peaks in the vocal tract impedance spectrum were usually smaller than those measured for the trumpet bore. In turn, they concluded that “[o]ver the range measured, none of the trumpeters showed systematic tuning of the resonances of the vocal tract” but conceded a possible invasive effect of placing the impedance head within the mouth of their participants, which “prevented them from playing the very highest notes of which they were normally capable” (abstract). Even the highest notes players managed to sound during the experiment (in the range from 1 kHz to 1.5 kHz) did not seem to require vocal tract tuning (p. 727); the same applied to pitch bending (p. 726). It is important to note the considerable variation in vocal tract resonance frequencies measured in this study, which suggests the use of a wide range of different vocal tract configurations over the playing range (p. 727; cf. findings of previous empirical research on brass instrument playing reviewed in section 2.5).

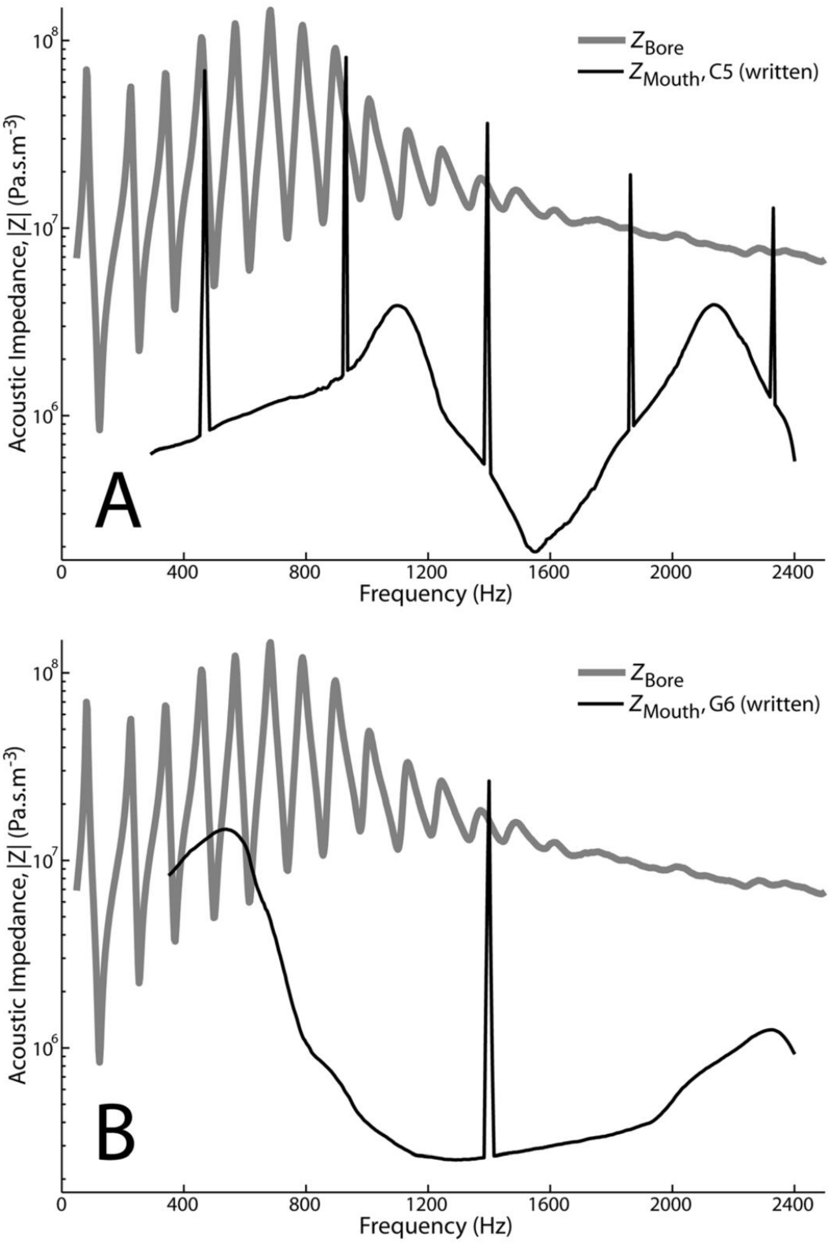


Figure 2.9: “In this semilog plot, Z_{Bore} [instrument impedance] and Z_{Mouth} [vocal tract impedance] are plotted for the notes written (A) C5 and (B) G6. In both cases, the note is played and Z_{Bore} is measured with no valves depressed. On the plots of Z_{Mouth} are superposed an artifact: The narrow peaks are the harmonics of the note being played.” Reprinted with permission from Chen, J.-M., Smith, J., & Wolfe, J. (2012). Do trumpet players tune resonances of the vocal tract? *The Journal of the Acoustical Society of America, 131*(1), 722-727. Copyright 2012, Acoustic Society of America.

Figure 2.9 on the previous page shows example impedance measurements in the mouth and bore for two different notes; it can easily be seen that the wide vocal tract impedance peaks (similar to speech formants) are very different for both notes (the sharp peaks are a measurement artifact). All players in the study reported raising their tongues to reach the very highest notes, which would appear to be consistent with the wide resonance peaks displayed in the figure (although they do not map directly onto speech vowel formants, the vocal tract configurations in A and B would probably be classified as different vowels in most languages, implying different underlying tongue positions). The authors speculate that raising the tongue, if not for vocal tract tuning, might facilitate high note playing by changing the magnitude or phase of vocal tract resonances (p. 727). We will see later on that indirect measurements of vocal tract tuning provide more explicit evidence for phase tuning of vocal tract resonances, and a resulting change in the magnitude of vocal tract impedance.

Similar findings were borne out in a study investigating the “[r]elationships between pressure, flow, lip motion, and upstream and downstream impedances for the trombone” (title of the paper by Boutin et al., 2015). In their conclusion, Boutin et al. (2015) state that trombonists did not tune their vocal tract for the low and medium pitch notes investigated in the study, attributing the lack of an effect to the “much lower” vocal tract impedance measured near the lips as compared to “that of the bore at playing frequencies” (p. 1208). When compared with readings from the study on trumpet playing, the measured vocal tract resonances did not vary as much for the different notes played. Figure 2.10 on the next page illustrates how the first vocal tract impedance peak was “always located between 200 and 325 Hz,” while measurements for the second resonance were “spread over a wider range […] between 513 and 985 Hz” (pp. 1199-1200). The authors interpret changes in the first tract resonance as mostly driven by changes in glottis opening (cf. section 2.3.2 below; the amplitude of the first impedance peak, however, also depends on the “geometry of the vocal tract”), and limited to a range of roughly 100 Hz, which may explain why players do not seem to vary the first tract resonance during trombone performance (p. 1200). The second tract resonance, however, could “presumably be modified by varying the position and shape of the tongue, as is done in speech to vary the resonances of the tract” (p. 1200). The observed values plotted in figure 2.10 show a split into two groups, with one centered around 650 Hz and another one around 900 Hz. For beginning trombone players (distinguished from amateurs and professionals), the values measured for the second tract resonance fell in the higher frequency group more often ("41% of notes played") than for advanced players ("18% of the notes"; p. 1200). Interpretation of these values based on the first speech vowel formant (F1, corresponding to the second vocal tract resonance peak; cf. figure 2.11 below) suggests the use of a lower tongue position by more proficient players.

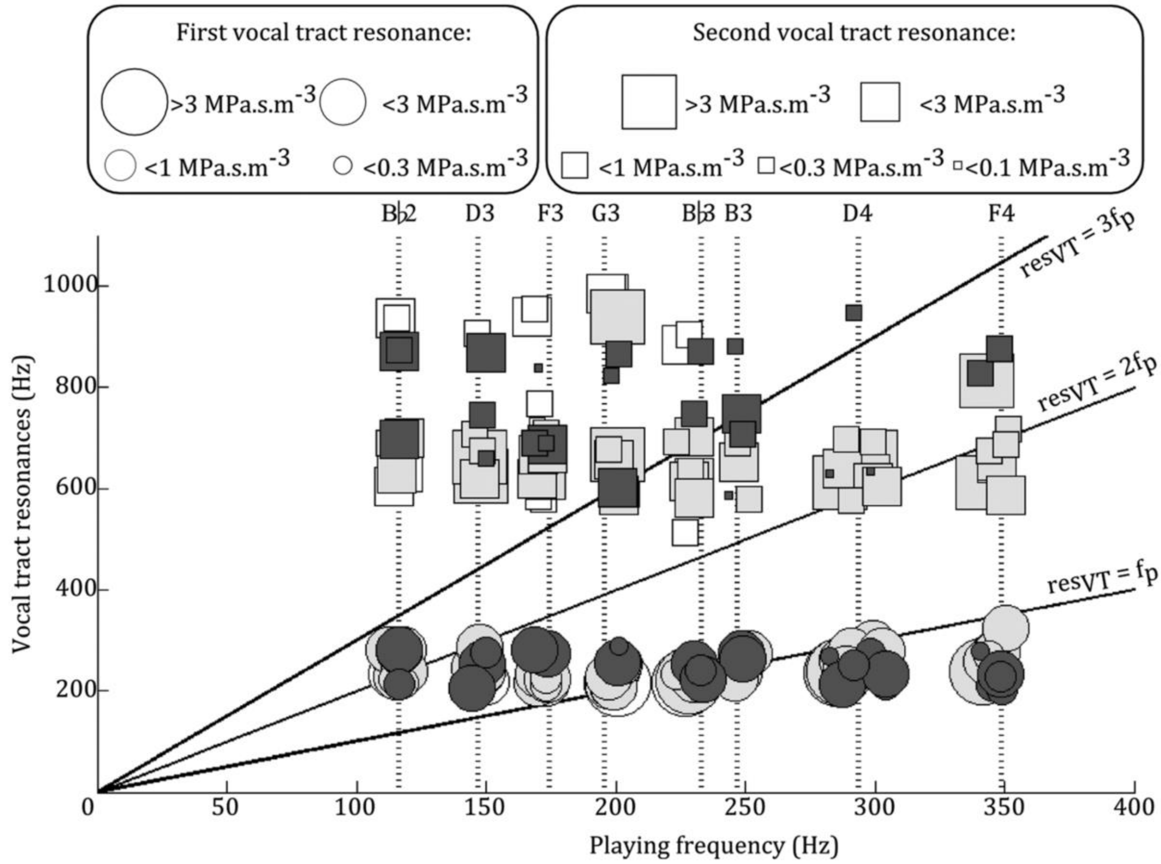


Figure 2.10: "Frequency of the first two maxima in the vocal tract impedance Z_{mouth} (resonances) compared with the playing frequency and the next two harmonics. The open symbols are data for beginners, the shaded symbols for amateurs and the darkest for professionals. The magnitudes of the first and second measured maxima of Z_{mouth} are binned in half decades, as shown in the legend and depicted by circles and squares, respectively. The three solid lines indicate when the vocal tract impedance peaks res_{VT} would be equal to the playing frequency f_p, and its harmonics $2f_p$ and $3f_p$. The dotted vertical lines indicate the nominal frequencies of the notes played." Reprinted with permission from Boutin, H., Fletcher, N., Smith, J., & Wolfe, J. (2015). Relationships between pressure, flow, lip motion, and upstream and downstream impedances for the trombone. *The Journal of the Acoustical Society of America, 137*(3), 1195-1209. Copyright 2015, Acoustic Society of America.

Before moving on to the findings gained from indirect measurements of vocal tract tuning, it is worth mentioning that the UNSW research group has also mentioned and, to a limited extent, investigated the possibility of vocal tract resonances influencing the timbre of wind instruments. It is likely that language influence would take such a form, provided that players from different language backgrounds produce the same pitches on their instruments. Although not determining or significantly affecting the frequency of the fundamental of a played note, such a "filtering effect, though smaller for most wind instruments than for voice," would admit the flow of acoustical energy into the instrument at some frequencies while inhibiting it at others (Wolfe, Garnier, & Smith, 2009, pp. 7-8). The effect has been applied as an advanced playing technique in a few compositions for trombone as a solo instrument (Erickson, 1969; Berio, 2006; cf. Wolfe, Garnier, & Smith. 2009, p. 8). Investigating such influence requires measuring vocal tract resonances throughout the whole playing frequency of a given instrument; the only study reporting such measurements for lip-reed instruments, according to Li et al. (2015), is Tarnopolsky et al. (2006) on the didgeridoo. The effect is "incontestable" regarding the timbre of the didgeridoo (Wolfe et al., 2003, p. 307) but much weaker on the trombone due to its higher impedance peaks and an additional formant introduced by the mouthpiece (cf. Wolfe et al., 2003, p. 310; Wolfe et al., 2013, p. 329).

2.3.1.2 Indirect measurement of vocal tract influence on brass instrument sound

A team of researchers at McGill University in Canada has conducted research on wind instrument and trombone playing using indirect measurements of vocal tract influence, based on the theoretical assumption specified in section 2.3.1. Their paper on trombone playing published in the *Journal of the Acoustical Society of America* (Fréour & Scavone, 2013; cf. Fréour, 2013; Fréour & Scavone, 2010) outlines a benefit of estimating vocal tract influence (indirectly) by measuring acoustic pressure, namely that it allows them to consider "both amplitude *and phase* of downstream and upstream system impedances at the frequency of interest" (p. 3888; emphasis added; note that a similar measurement was included in Boutin et al.'s, 2015 paper). This data can be used to estimate the nature of the lip vibrating mechanism employed throughout the different registers of the instrument and forms an important part of the study's findings. Overall, the results of this study agree with the papers reviewed above regarding vocal tract tuning in the normal playing range of the trumpet and trombone. However, for two notes in the trombone's extreme high register, the authors' observations suggest the optional use of two different strategies of vocal tract tuning: one involves producing a large upstream (vocal tract) impedance amplitude (compared to the bore impedance) at the frequency of a note's fundamental, thus "overriding the effect of the trombone impedance" (p. 3897); the other strategy is to carefully tune the phase of the upstream impedance at the fundamental "which might better support oscillations near a mechanical resonance of the lips" (p. 3894). All participants were observed to follow a trend: a reduction of phase differences between upstream and downstream pressure flows coincided with rising pitch, indicating the dominance of downstream coupling in the low register and constructive phase tuning in the higher range (p. 3891). Additionally, an increase in upstream impedance (estimated from the pressure differential between the two cavities) was observed when transitioning between slurred notes (pp. 3896-3897), while an increase in loudness seemed to lead to a decrease in vocal-tract support that could be related to "non-linear interactions between harmonics" at extreme dynamics (p. 3897).

In a follow-up paper (Fréour et al., 2015), the authors collaborated with researchers from the Institut de Recherche et Coordination Acoustique/Musique (IRCAM) in France to further investigate the phase tuning of vocal tract resonances in brass playing, employing an artificial player system and numerical simulations. The artificial player system used for their "in-vitro" investigation (Hélie, Lopes, & Caussé, 2012) allows the adjustments of the "amplitude and phase of the pressure at the input of the mouth cavity […] relative to the pressure at the input of the downstream air-column" (p. 258)[4]. The system's upstream cavity (artificial mouth) has a resonance of around 400 Hz which is comparable to the first resonance peak of the vocal tract found in studies of human players, with "[t]he amplitude of this impedance peak being less than half the amplitude of the closest downstream resonance" (p. 258). Focusing on "the highest tones that can be produced by the artificial player without active upstream feedback" based on previous research that showed vocal tract influence to be most pronounced in this range, the phase of the upstream impedance was modified via a linear phase sweep over a range of 240 degrees in sixty seconds, all while holding

[4] Although these changes were only effected at the fundamental frequency of the investigated notes, the authors add that it would be "theoretically possible to impose an upstream feedback with several harmonics of the fundamental frequency," thus making it possible to study the effect of vocal tract tuning on the timbre of brass instruments (Fréour et al., 2015, p. 258), comparable to the study by Li et al. (2015) investigating saxophone performance.

upstream acoustic pressure constant, “regardless of variations in the” resulting “downstream acoustic pressure” (p. 260). “[A] potential optimum tuning of the system” could be observed at a particular value of the relative phase difference where the downstream acoustical feedback (on the lips) not only reached a maximum, but resulted in the maximization of lip motion and hence “a larger acoustic flow into the instrument” (p. 262). Furthermore, significant variations of playing frequency were observed in the range of 418 to 408 Hz for the note Ab4, and changes in the amount of upstream acoustic energy injected into the system were significantly correlated with such variations (pp. 261-262). Numerical simulations of “a simple outward striking” and a more complex two-dimensional lip model modified by “a disturbing upstream signal representing the effect of an upstream resonance” similarly showed that variations of the phase of an upstream resonance relative to downstream impedance lead to variations in playing frequency (p. 268).

2.3.2 The role of the glottis

Recent research on the role of subglottal resonances in speech production has demonstrated their relevance for speech acoustics (Stevens, 1989; Hanson & Stevens, 1995, Stevens, 2002; Chi & Sonderegger, 2007; Jung, 2009; Lulich, 2010; Lulich et al., 2011), revealing a pattern of upstream influence on downstream oscillations not unlike the effect of vocal tract influence during brass instrument sound production. Even though it does not act as the sound-generating valve during brass playing, the glottis nevertheless plays an important role by regulating the contribution of the subglottal cavities to upstream impedance. According to Wolfe, Garnier, & Smith (2009), “the size of the glottis affects both the frequency and the magnitude of peaks” interacting with the oscillating airstream within the instrument (p. 10). Seemingly contrary to intuition, experienced brass players (and players of other wind instruments; Mukai, 1992) use a “much smaller glottal aperture” during playing than beginners (Wolfe et al., 2015, p. 3, discussing findings from Mukai, 1992; cf. Dejonckere et al., 1983; King, Ashby, & Nelson, 1989; Rydell et al., 1996). It appears that restricting glottal aperture during playing is a learned behavior that develops with proficiency, but other changes in glottal configuration during playing may be more coarticulatory in nature. Bailey (1989) used laryngoscopy to observe the vocal fold position in trumpet players and found that changes were rather small and moreover “self-adjusting or involuntary” (p. 105; cf. section 2.5.1 below). Electroglottographic data of French Horn players by Dejonckere et al. (1983) corroborates this finding; they registered a signal resembling “the graphic of a breathy sound uttered by a male subject in falsetto voice” which they interpret as evidence “that during horn playing the vocal chords maintain a position relatively close to each other, thus narrowing the respiratory airway” (p. 33). The expiratory airflow would thus be “able to stimulate the glottis oscillator” (p. 33), a behavior also observed by Wolfe & Smith (2008) during didgeridoo performance (documented in Wolfe, Garnier, & Smith, 2009, p. 11).

Why do advanced wind players seem to restrict glottis opening? Wolfe et al. (2015) believe the motivation behind this could be to “allow fine control of mouth pressure” and “providing a higher reflection coefficient for acoustic waves in the vocal tract” (p. 3). *In-vivo* measurements of vocal tract influence on trombone playing by Boutin et al. (2015) reviewed in the previous section seem to support this assumption: in this study an upstream (vocal tract) impedance maximum was consistently measured around 200 Hz, which the authors assume players can only vary over “a range that is less than 100 Hz [and] which may partly explain” the absence of any evidence for vocal tract tuning (p. 1200). As stated in an earlier paper (Chen, Smith, & Wolfe, 2012), this group of researchers believes the influence of tongue position to be relatively weak for the first vocal tract resonance, which depends rather strongly on the area of the glottis, while “the shape of the tongue, and especially the position of a constriction between the tongue and the roof of the mouth, can vary the frequencies of the second and third resonances considerably” (corresponding to F1 and F2 in speech; Chen, Smith, & Wolfe, 2012, pp. 724-725). Evidence for the considerable influence of glottis opening in the low frequency domain comes from two papers by Hanna, Smith and Wolfe (2012 & 2016; cf. Titze, 2008; cf. Henrich et al., 2005 for the effect of glottal open quotient on singing). Figure 2.11 on the next page shows “two measurements of a vocal tract, one during exhalation and one with the glottis closed” (Wolfe et al., 2015, p. 11). Both impedance curves closely resemble one another in the high frequency domain, but differ in behavior at lower frequencies, “which presumably is important for transients [tone buildup at beginning of notes] and possibly for low notes” (Wolfe et al., 2015, p. 11). Actual vocal tract impedance during brass playing should pattern somewhere in between the impedance functions plotted in figure 2.11 (based on findings of glottal aperture from empirical research cited above) and could vary somewhat among players (cf. Wolfe et al., 2015, p. 1200). Nonetheless, the combined research findings suggest that the influence of the first vocal tract resonance on brass playing should be limited due to its restricted adjustability.

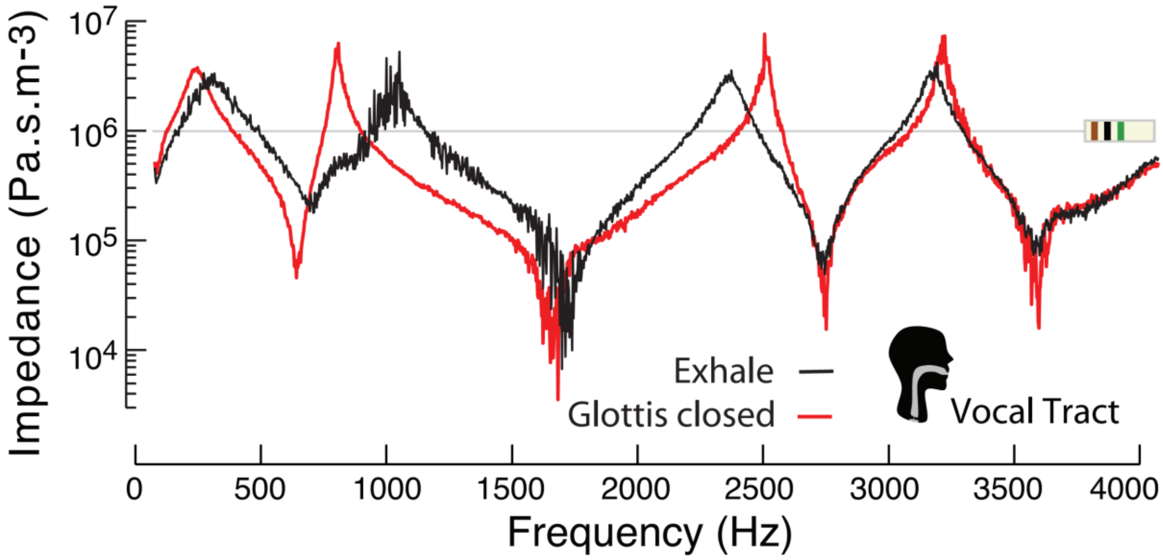


Figure 2.11: The impedance spectra of a vocal tract measured at the lips: glottis closed (red) and exhaling (black); note that the minima around 650 and 1700 Hz correspond to the formants of [ə] produced by this simple vocal tract configuration (cf. Hanna, Smith, & Wolfe, 2012, p. 3). Reproduced with permission from Wolfe, Fletcher, & Smith, 2015, p. 11; first printed in Hanna, Smith, & Wolfe, 2012.

2.3.3 Summary

The literature reviewed above provides no evidence for the systematic tuning of vocal tract resonances during brass instrument playing (and trombone playing, specifically), and the observed individual variation suggests that different players may employ different strategies to optimize tone production. Despite the fact that there is currently no conclusive evidence, the underlying physics, however, do suggest that vocal tract influence on brass instrument sound should be possible; the world-leading UNSW research team certainly seems to have considered this possibility (cf. Wolfe et al., 2003, p. 310; Wolfe et al., 2013, p. 329). Unfortunately, the effect of vocal tract configuration on brass instrument timbre, specifically, has not yet been the focus of investigation (aside from the didgeridoo, cf. section 2.3.1.1 above).

Another assumption concerns the extreme registers of brass instrument playing. I interpret the limited findings in the literature to indicate that the higher one aspires to play, the more important the involvement of the vocal tract becomes, along with its tuning to the fundamental frequency of the played note. If this assumption is correct, it could also mean that lower notes can be produced by employing a variety of vocal tract shapes with an acoustic effect on timbre, rather than determining whether a particular note is sounded. This could then be the area where language influence manifests itself most clearly: with fewer constraints arising from acoustical considerations there would be more space for the vocal tract configurations of one's L1 to impact the produced sound. Note that I explicitly do not exclude effects of differences in individual vocal tract morphology (cf. section 3.3.4) on brass instrument timbre, which could not be addressed in this study.

In addition to its plausibility in terms of the underlying physics, the assumption that vocal tract configuration can influence brass playing receives considerable support from hundreds of years of writing on brass pedagogy; this literature is reviewed in the following section.

2.4 Brass pedagogy

2.4.1 The use of speech syllables in brass pedagogy

This section reviews the use of speech syllables in the history of brass instrument teaching, starting with a method book on recorder playing published in 1535, and the popularity of doing so continues into the present day, even after this notion was questioned by early empirical research (cf. Irvine, 2003; Schlafer, 2006 for the bassoon). The concept is most frequently applied to facilitate playing throughout the various registers of the instrument, where the use of different vowel tongue shapes is thought to help regulate the required airflow. The application of different vowel tongue shapes represents an interesting possibility in terms of language influence on brass playing, particularly in light of the fact that brass pedagogy has almost exclusively used 'cardinal' or orthographic vowels, thus neglecting a significant portion of the vowel tongue positions occurring in natural speech.

2.4.1.1 Pedagogical writings on brass playing predating the modern era

Even prior to Dalla Casa's (1584/1970) treatise on the Renaissance cornetto[5] mentioned in the introduction, a number of authors (Ganassi dal Fontago, 1535/1956;

[5] This instrument is not to be confused with the modern cornet employed mainly in British-style brass bands. It is rather a finger-hole trumpet with a small cup mouthpiece now only played by a small number of musicians specializing in Early Music.

Agricola, 1545/1896; Cardano, 1546/1973) included syllables illustrating the use of the tongue in their treatises on wind instrument playing, although none of these sources referred to brass playing explicitly (cf. Heyne, 2015).

Early Italian sources largely agree on the recommended syllables (consonant/vowel combinations) to use during wind instrument playing; see Tables 2.1a-c below for a representative sample. The most commonly specified vowel is orthographic <e>, although Fantini (1638/1978) uses all five orthographically possible vowels in Italian (cf. footnote on pages 40-41 regarding Italian (historical) pronunciation). A common theme is the differentiation of three 'original tonguings', one of which is the so-called 'reversed' tongue ("lingua reversa"), described in the following passage from Dalla Casa:

> Reversed tonguing, being the principal one among the three tonguings, will be dealt with first, since it resembles the gorgia more than the others. It is [thus] called gorgia tonguing. This tonguing is very fast and is difficult to restrain, its striking point is at the palate, and it is pronounced in three ways: ler ler ler ler, der ler [der ler], ter ler ter ler. The first pronunciation is sweet, the second is intermediate, and the third, being a more pointed tonguing, is harsher than the others (1584/1970, p. iv; translation by Tarr & Dickey, 2007, p. 55).

This description of the kind of tonguing to be used when performing diminutions (fast added notes/embellishments) attests to the high regard held by Renaissance instrumentalists for the human voice as an ideal to be imitated. In fact, the alternative term given by Dalla Casa, 'lingua gorgia', means "throat tonguing", a reference to the fast articulations produced by the singers of the time (cf. Tarr & Dickey, 2007, p. 20). For the reader familiar only with contemporary performance technique, it is important to understand that, contrary to present-day practice, unequal tonguing was an ideal of 16th-century music to be employed in passages featuring large numbers of notes executed in quick succession. According to Tarr and Dickey (2007), "[t]he application of a compound tonguing" which employs different places of articulation thus "provides an ideal technical means of realizing this theoretically sanctioned inequality" (p. 20). Furthermore, individual differences in the ability of performing fast articulations received some early attention in Bendinelli's trumpet method:

> Moreover, one can never write down everything about tonguing since different people use their tongues differently in playing. Some [players] are by nature stutterers, [a fact] which makes it difficult for them to play using compound tonguing. These can learn to play with single tonguing, which will work very well for military signals… (Bendinelli 1614/1975, p. i^{v}; as cited in Tarr & Dickey, 2007, p. 81).

Only one these early sources distinguishes articulations for specific instruments; Bismantova (1677/1978) specifies <te> to be used when playing the cornetto, while he gives <de> for the recorder (Tarr & Dickey, 2007, p. 24).[6]

[6] It might also be relevant to know that, prior to the beginning of the nineteenth century, “articulation was seldom indicated in the music,” while in modern practice it “is usually notated by means of such symbols as slur marks, dots, wedges, accent marks, etc.” (Tarr & Dickey, 2007, p. 11). For an overview of the limited existing acoustical research on articulation and its psychoacoustics, see Spiegelberg (2002).

Source	Intended instrument/s	Articulation syllables used	Recommended vowels	Comments
Ganassi, (1535/1956, 1535/1969) – Opera intitulata Fontegara; information from Tarr & Dickey, 2007, pp. 18-19, 41-45)	recorder but Ganassi is thought to have played cornetto and sackbut (renaissance trombone) as well	three 'original tonguings': te che te che te che (/teketeketeketeke/) tere tere tere te (/tere tere tere te/) lere lere lere le (/lere lere lere le/)	main vowel <e>, but also mentions variants: "Tacha teche tichi tocho tuchu" with the following comment: "Note how I proceed with the vowels, in order that you may discover which syllable or letter Nature has granted you [the ability] to pronounce with greater speed" (Ganassi dal Fontago, 1535/1969, p. 5v, as cited in Tarr & Dickey, 2007, p. 43)	"Ganassi's articulations bear striking similarities to those in a series of later Italian sources, suggesting that his *Fontegara* was highly influential in the area of the Venetian Republic and even beyond" (Tarr & Dickey, 2007, p. 19).
Cardano (1546/1973) – De musica (treatise written in Latin); information from Tarr & Dickey, 2007, pp. 17, 50-52)	not specified (Cardano, "a physican and philosopher, was an amateur recorder player" (Tarr & Dickey, 2007, p. 17))	lere, theche / there, thara (note that the second syllable is spelled "theche" in the facsimile but incorrectly copied by Tarr & Dickey, 2007, pp. 51, 24) as "teche")	uses <a> besides the more common <e>	talks about 4 different kinds of tonguings, one of which is 'motu' (motion), for which he gives the descriptions "stretched out" and "bent back," which could refer to the use of tongue to modify tone quality and would be "unique among the sources known to the present authors" (Tarr & Dickey, 2007, p. 52)
Dalla Casa (1584/1970) – Il vero modo di diminuir; information from Tarr & Dickey, 2007, pp. 19, 53-60)	Renaissance cornett(o)	ler ler ler ler / der ler [der ler] / ter ler ter ler tere tere tere terete teche teche teche techete te te te te / de de de de	uses <e> exclusively	prints syllables underneath notes for some of the practical exercises included in treatise; see quote about 'reversed' tonguing above
Fantini (1638/1978) – Modo per imparare a sonare di tromba; information from Tarr & Dickey, 2007, pp. 18, 98-99)	trumpet	le ra le ra li ru li / ti ri ti ri ti ri di te ghe te ghe te ghe di ta te ta ta ti ta ta / la le ra la la la la lal de ra de ra de ra / la ra le ra la ra la te re te re te ghe te ghe di tia tia da -dia dia da (tia and dia represent two notes each)	all Italian (orthographic) vowels; offers two or three different options for the same sample passages	

Table 2.1a: Representative sample of articulation syllables used in early Italian sources.[7]

[7] Historically, two main varieties of Italian have been distinguished, which differ in phonological status regarding the realization of the mid vowels. While [e], [ɛ] and [o], [ɔ] are used interchangeably in Northern Italian (representing allophones), they are lexically distinctive in the Southern variety (Clivio, 1996). We can presume that, of the authors

Source	Intended instrument/s	Articulation syllables used	Recommended vowels
Arbeau (1588/1967, 1588/1972) - Orchésographie; information from Tarr & Dickey, 2007, pp. 18, 98-99)	dance manual	two manners of tonguing on the flute: using the syllable te – té té té or tere tere tere; 'rolled' tonguing – relé relé rel (accent suggests stress on second syllable, similar to 'lingua reversa' of the Italian school)	clearly differentiates two realizations of <e> via accent markings: <é> [e] and <e> [ɛ] or [ə]; see footnote below
Mersenne (1636/1957, 1636/1965) - Harmonie universelle V; information from Tarr & Dickey, 2007, pp. 17, 94-99)	not specified (Mersenne's work represents a theoretical contribution, he seems to have gathered his technical information from practicing musicians)	Tarare, Tarararare	uses <a> almost exclusively
Laquemant & Chrestien (1666/1998) – Recueil de pièces de viole en musique et en tablature; information from Tarr & Dickey, 2007, pp. 100-101)	signals for hunting horn contained in manuscript of unaccompanied solo viol music	ti ta, ha (for tied notes only)	uses ti consistently for high notes (local context), ta for low notes

Table 2.1b: Representative sample of articulation syllables used in early French sources.[8]

included in table 2.1a, Dalla Casa, and Cardano (although he wrote his treatise in Latin) would have spoken Northern Italian. Although the standard spelling system of Italian nowadays uses accents to distinguish the mid-vowels in final position (cf. Clivio, 1996, pp. 210-211), such a standard did not exist in the sixteenth and seventeeth centuries, and the nonsensical status of articulation syllables per se precludes a sound linguistically-grounded determination of their vowel qualities. No distinction can thus be made based on the origin of the authors; they all, however, stress the importance of ease of articulation, providing a hint that a higher tongue position ([e] instead of [ɛ], [o] instead of [ɔ]) may have been intended, especially at fast tempi.

[8] The situation regarding the front mid vowels in Renaissance French (1450-1600) is similar to Italian; <e> was pronounced [ɛ] in open syllables and [e] in checked syllables (including a consonant coda). However, in (stressed) word-final environments where consonants gradually disappeared (e.g. *bontét* becoming *bonté*), [e] in final position could result. Furthermore, unstressed <e> became increasingly reduced throughout the history of French, moving towards [ə] in Renaissance French before disappearing altogether in final unstressed position in standard modern French (cf. Price, 2005, p. 81). In contrast to modern spelling (and modern editions of historical sources), these pronunciations were not usually differentiated by the use of the acute accent in historical publications (Taylor, 1996). Luckily, Arbeau (1588/1967, 1588/1972) provides us with accent markings disambiguating the intended realization of vowels in his syllables, apart from the final vowel in "tere" which could be realized as either [ɛ] or [ə]. Both <a> and <i> are unambiguous as their realization in Renaissance French would have been identical to modern French: [a] and [i].]

Source	Intended instrument/s	Articulation syllables used	Recommended vowels	Comments
Agricola (1545/1896) – Musica instrumentalis deudsch; information from Tarr & Dickey, 2007, pp. 17, 46-49)	not specified (Agricola was a music theorist and recommended seeking specialist instruction, in addition to reading his book: “Now, if you further make the decision; To learn the art of pure division And place the mordants quite correctly; Then to a teacher go directly – For I cannot speak of these; In such a place as this with ease” (no p., cited from Tarr & Dickey, 2007, p. 47).	de de de de de - di ri di ri di ri de	<e>, <i>	
Altenburg (1795/1974 & 1795/1993) – Versuch einer Anleitung zur heroisch-musikalischen Trompeter- und Pauker-Kunst... [Essay on an Introduction to the Heroic and Musical Trumpeters' and Kettledrummers' Art ...]; information from Tarr, 1974, pp. 91-93)	courtly trumpet and kettledrum playing (it is important to note that at the time of Altenburg's writing, the Golden age of the baroque trumpet was already over, "since trumpeters had already been nearly completely deprived of their courtly basis of existence" (Tarr, 1988, p. 9).	single: ri-ti-ri-ti-ton, ki-ti-ki-ti-ki-ton; double: ti-ri-ti-ri-ton, ti-ki-ti-ki-ton; Altenburg says these are only used in the low register but does not provide other syllables to be used in the other registers.	<i>, <o>	also talks about 'huffing' (German “Überschlagsbindung”), for which he gives the syllables to-ho-to (huffing happens on ho syllable)

Table 2.1c: Representative sample of articulation syllables used in early German sources.[9]

[9] In German, the pronunciation of the front mid vowel <e> patterns the other way around compared to Italian and French; it is pronounced as a long vowel [eː] in open syllables and as [ɛ] in checked syllables. This situation was already established by the time of Agricola’s writing, in a period of the German language referred to as Middle High German or New High German (1150-1350; Frenzel, 1996). Similar to Renaissance French, <e> in final syllables was probably already realized as the (high) schwa [ə] of modern German (Penzl, 1975). The other two German vowels used by Altenburg, <i> and <o>, were realized as [iː] and [oː], or [ɪ] and [ɔ] in open and checked syllables, respectively. A final uncertainty regarding the pronunciation of articulation syllables in the French and German treatises concerns the question of whether their articulation would change when applied in sequence or close proximity. Should the vowel in the final syllable of Altenburg’s “de de de de de” differ in quality from the preceding ones?

Possibly the earliest account of playing differences among brass musicians from different language and/or cultural backgrounds is a scene from a novel (1683) attributed to Daniel Speer (1636-1707), who was also a composer of brass music.

> The hero of Speer's tale, a trumpeter, was led to a new encampment in Hungary and listened uncomfortably as two Gypsy trumpeters tried 'obscenely bravely' to play German and Polish pieces. When urged himself to play, he played a 'dikedank and dikedikedank in the German manner, which was so strange to these people that they all stood still, giving audience to my [sic] German and, then, Polish 'Trippel-Liedchen' …' (Speer, 1683/1973, p. 153; as cited in Tarr & Dickey, 2007, p. 25).

As pointed out by Tarr & Dickey (2007, p. 25), a clear connection can be made between this literary quote and the articulation syllables[10] later published by Altenburg (1795/1974 & 1795/1993; see table 2.1c).

2.4.1.2 Pedagogical writing on brass playing published in the last fifty years

Modern brass pedagogy has continued to use speech syllables even after the expedience of such a practice was questioned by early empirical studies using x-ray imaging (see section 2.5.1 below; cf. Irvine, 2003). This section provides some information about the more recent historical background and describes a small number of important pedagogical approaches of the last fifty years.

Loubriel (2011b; based on McCann, 1989) reviewed the main pedagogical trends in trumpet playing prevalent during the early twentieth century. Among these are the so-called 'tongue arch techniques' advocating alterations of tongue shape to facilitate lip slurs, represented (among other method books published in 1922 and 1936) by Schlossberg's *Daily Drills and Technical Studies for the Trumpet* (Schlossberg, 1937; Loubriel, 2011b, p. 3), which continues to be used today (cf. Larson, n. d.; Hauf, 2007; Quinque, 1982; in addition to sources reviewed below). Even though I cannot personally recall being told to use different vowel tongue shapes in different registers of the instrument, I know from conversations with other brass players that this approach is still frequently taught in Germany, and probably many other places. Note that the application of speech syllables is not necessarily dependent on their use in

[10] Note that while I have chosen to use the term articulation syllable in this section to stay consistent with its use in the book by Tarr and Dickey (2007), I understand it to be synonymous with the term speech syllable used throughout this thesis.

method books; many teachers use or choose not to advocate them irrespective of their appearance in such publications. There is also a contrary approach that recommends minimizing or eliminating changes in tongue position altogether; as far as I know, neither side of the dispute has seriously considered empirical evidence for or against their case. There have been, however, a limited number of attempts at introducing scientific inquiry into brass teaching (cf. Loubriel, 2011b, p. 6), mainly by educators carrying out their own research or collaborating with experts in the field (see section 2.5 below; cf. Weast, 1963; Carter, 1966/69; Haynie, 1969). More recently this situation has improved; Steenstrup (2007), to mention the foremost example, discusses a great range of empirical research available at the time of his book's writing. In the paragraphs below I will briefly review a necessarily limited selection of pedagogical publications making use of the tongue arch technique.

Fred Fox (b. 1914) is an American French horn player (he retired from performing in 1969; Playing under Toscanini: An alumnus remembers, 2014) whose book *Essentials of Brass Playing* (1982) "is considered the bible of brass technique" (Alimurung, 2014). He recommends the use of vowels to work out how one can "produce a good tone on any brass instrument simply by adjusting the tongue position to the needs of that particular instrument" (p. 12). Though Fox assigns various vowels or syllables to different registers ("Aw", "E", "Ee" in ascending order, with "Iss" recommended for the "highest notes"; p. 16), he encourages players to experiment to find the sound they want:

> Play a note in the middle register of your instrument using an extreme "ee" (as in the word "see") tongue position. The note will sound thin. Play the same note again using an extreme "aw" (as in the words "law") tongue position. The note will now sound dull, like a foghorn. Play the note a third time with the tongue midway between both extremes. The sound is now more satisfactory because you have both the highs and lows present in the tone. Some players may prefer a few more highs or lows […] but there should never be an extreme imbalance in either direction (p. 11).

By regarding vowel tongue positioning as a continuum and suggesting instrument differences, Fox's approach represents one of the more flexible accounts among the tongue arch techniques[11].

The Encyclopedia of the Pivot System for all cupped mouthpiece brass instruments: A scientific text by Doctor Donald S. Reinhardt (Reinhardt, 1973; first published in 1942) asserts a strong claim of scientific objectivity in its title[12]. Contrary to what an academic might expect, however, the book does not present any empirical evidence, although Reinhardt provides remarkably detailed accounts of the (visible) changes occurring during the performance of brass players with a focus on reoccurring individual patterns. This wealth of description is unequaled in any other teaching method I know, and feeds into the identification of two distinct 'pivot' types, four standard plus five subtypes of embouchures, and eight different tongue types. The use of the term pivot here refers to "a vertical movement of the mouthpiece and lips on the face of the front teeth" (Turnbull, 2001, p. 5), which seems to recur (for most players) around the same pitch within the mid-range of all brass instruments[13]. The existence of such a phenomenon is now widely accepted (albeit with some confusion regarding the exact nature of the phenomenon, cf. Turnbull 2001, p. 5), as is the classification of brass players' embouchures into 'downstream' and 'upstream' types[14] that relates to whether the upper or lower lip are dominant in creating the lip vibrations initiating brass sound. As far as I know, no other brass pedagogues have postulated the existence of such a large number of anatomically different tongue types, though researchers in speech production have certainly considered the influence of palate

[11] See also his concluding sentence: "The vowel position varies in the different registers of one instrument. Furthermore, the overall vowel sound varies from instrument to instrument. No one tongue position can be correct from the lowest to the highest note on any particular instrument, or from soft to loud on the same note. The ratio of highs and lows must remain constant throughout the entire range of the instrument. This can be accomplished only by listening and constantly adjusting the vowel sound" (Fox, 1982, p. 18).

[12] The preface makes it clear what the author means by the use of the adjective 'scientific':
"The PIVOT SYSTEM is a scientific, practical, proven method of producing the utmost in range, power, endurance and flexibility on the trumpet, trombone and all other cupped-mouthpiece brass instruments. It was originated not only through forty years of research and experimentation in practical playing, teaching, writing and lecturing to many thousands of professionals, semi-professionals, supervisors, teachers, students, etc., but also through designing and producing personalized mouthpieces and being consultant of instrument design for several leading manufacturers of brass instruments" (Reinhardt 1973, p. F).

[13] It is possible that this movement is related to changes in the lip vibrating mechanisms in different registers (discussed in section 2.3.1.2 above) but I am not aware of any research that has investigated this hypothesis.

[14] This use of the term 'upstream' is not to be confused with its use in musical acoustics where it refers to the existence of a resonator upstream from the source.

shape and other aspects of vocal tract morphology on tongue kinematics (Stone & Vatikiotis-Bateson, 1995; Fuchs, Winkler, & Perrier, 2008; Brunner, Fuchs, & Perrier, 2009; Thibeault et al., 2011; cf. section 3.3.4 below). Concerning the "tongue-arch level" (which he also relates to "breath focus"), Reinhardt implicates it as "a prime controlling factor for: 1. range; 2. endurance; 3. flexibility; 4. intonation; 5. tonal timbre; 6. mouthpiece pressure; [and] 7. centering the sound" (p. 71). Though he gives vowel recommendations for the various registers, "AAA for the lower register – OOO for the middle register – EEE for the upper register" (p. 95), he acknowledges individual differences:

> Since no two mouth cavities are identical, no two performers will enunciate the syllables TAAA or DAAA (for the lower register) – TOOO or DOOO (for the middle register) – and TEEE or DEEE (for the upper register) in the same manner or in the same position in their mouths. Therefore, individual differences are being given their usual careful consideration even though identical syllables are used (p. 83).

Overall, Reinhardt's careful descriptions can be regarded as a major step towards empirically informed practice in brass teaching[15]; unfortunately, his theories were often misunderstood and met with uninformed criticism by the brass world (cf. Turnbull, 2001).

Arnold Jacobs (1915-1998), longtime principal tuba player of the Chicago Symphony Orchestra, was possibly the most highly regarded brass teacher of the twentieth century; in an era of increasing specialization, he taught players of all kinds of brass instruments and even woodwind players and singers (cf. Stewart, 1987; Frederiksen, 2006; Loubriel, 2011a&b). He believed that the color of tone could be changed by using vowels (Frederiksen, 2006, p. 128), but like Fox, stressed that the exact shape to be used was to be determined by great sound: "Ultimately, a player will play anywhere in between the 'AH' and the 'EEH' tongue positions" (Loubriel, 2011b, p.

[15] Reinhardt also claims that "[u]pstream types generally sound more brilliant, but less resonant, than the downstream types. This is one of the reasons why the performers in the symphonic field are rarely of the upstream variety" (p. 224), and "[d]ownstream types generally sound more resonant, but less brilliant, than the upstream performers" (p. 225). I do not believe this has been tested empirically; a different "breath focus" (as Reinhardt calls it; p. 71) could necessitate a change in tongue position which might in turn influence timbre (cf. Heyne & Derrick, 2015a).

78)[16]. Another parallel concerns Jacobs' regard for different tongue morphologies and how these might affect brass playing:

> Some people's tongues take up most of the oral cavity. There are others with huge tonsils, some have moderate tonsils and a fairly big tongue. I frequently find crowded conditions to the point where even relaxed and properly used oral equipment does not work right simply because it is taking up too much space (Frederiksen, 2006, p. 126).

Unlike the other pedagogical approaches reviewed here, Jacobs was aware in his teaching of the difficulties involved with trying to consciously control a muscle as complex as the tongue, recommending the use of enunciation exercises to train tongue movements (Frederiksen, 2006, p. 127). This did not guard him against making incorrect assumptions, however, such as recommending "[t]o experience an open airway, say 'ah, oh, ooh'"; it has been shown that "ah" (/ɑ/) indeed opens up the oral cavity but significantly constricts the pharynx (cf. Baer et al., 1991).

Finally, a short discussion of American French horn player Eli Epstein's teaching philosophy is warranted here due to his involvement with some recent research using Magnetic Resonance Imaging (MRI), summarized in section 2.5.3 below. He proposes using a concept called 'Finger breathing' (first suggested by flutist Keith Underwood) to select the correct vowel tongue shapes during brass playing. This involves holding a flat hand vertically in front of the opened mouth while touching the lips with the base of the forefinger in a midsagittal orientation. Breathing in deeply in this configuration produces a strong turbulent sound where the airstream hits the hand, and the idea is to change the perceived pitch of this noise by adjusting tongue position (Epstein, 2016). The underlying assumption is again the use of a 'tongue arch' employing different vowel tongue shapes to facilitate brass playing. Note, that in a different video with Peter Iltis (Epstein & Iltis, 2016b), Epstein presents midsagittal MRI images of players producing notes and sustained vowels as evidence for the similarity of these tongue gestures while they are actually very different (cf. footnotes in section 2.5.3 below).

[16] Jacobs seems to have directly pointed out the intended physical consequences of using different vowel tongue shapes to his students, as this quote from a teaching situation indicates: "Later you can find the tongue level and the tone production you want. I want you to be able to function with a closed or open oral cavity. … it just has to sound great" (Loubriel, 2011a, p. 78). He lists the following good vowels: "AH, OH, or OOH" (Loubriel, 2011a, p. 76).

The use of tongue arch techniques has also led to a lot of confusion when method books have been used by teachers and students who do not share the same native language and orthography as the author (cf. Zsaisits, 2012, pp. 47, 64). For example, the articulation syllables “tukutu” printed in Arban’s famous method for trumpet (Arban, 1936) are frequently interpreted as /tukutu/ in Germany and /tʌkʌtʌ/ in US, while the author probably had something more like /t̪ykyt̪y/ in mind; the multilingual edition of Quinque’s *asa technik* (1982), however, provides explicit directions for non-German speakers on how to achieve the desired vocal tract configurations for <ö> ([ø, œ]) and <ü> ([y, ʏ]):

> ö (cf. French eu) - Say the vowel ay (as in day); now, keeping the tongue in the same position, round the lips in to the position for saying the vowel oo; ü (cf. French u) - Say the vowel ee; keeping the tongue in the same position, round the lips into the position for oo (Quinque, 1982, p. 13).

In a 2009 article for the *Journal of the International Trumpet Guild*, American trumpeter Jason Dovel also cautions against the frequent use of diphthongs instead of ‘pure vowels’ when demonstrating vowel tongue positions to be used during playing (Dovel, 2009).

2.4.1.3 Summary of speech syllables used in brass pedagogy

Throughout the history of brass playing and teaching, various pedagogues have used speech syllables to illustrate what students should do with the tongue while playing their instruments.

The earliest available sources listed the following vowels, classified here by the (presumed) first language of the authors:

Italian: /e~ɛ/ but also /a, i, o~ɔ, u/

French: /e~ɛ/ but also /a, i/

German: /e/, i/ but also /o~ɔ/

Method books and teaching philosophies circulated since 1942 include the following combinations of vowels: /ɑ~a, ɛ~e, i/ or /ɑ~a, o, u/.

These vowels represent a pattern of tongue raising either along the front or back sides of the traditional vowel trapezoid. While proponents of the so-called tongue arch techniques advocate strict adherence to these shapes when ascending throughout the pitch range of a given instrument, authors interpreting this concept more liberally regard the suggested tongue shapes rather as an approximation.

It is striking that one cannot find any evidence of the suggested use of central vowel tongue positions, such as schwa (/ə/) in the literature; this could be due to a lack of representation in standard orthography. Overall, even in languages with complex vowel systems, the recommendation to use cardinal or cardinal-like vowel qualities predominates.

2.4.2 A personal account of recent developments in the world of brass playing

Readers unfamiliar with the profession of classical/orchestral music performance may not be aware that there exists a prominent concept referred to as national schools or styles of playing. This concept relates to differences in playing style co-occurring with national borders and/or language boundaries, and while the concept provides an important reason to presume that there should some kind of language influence on brass playing on the one hand, it represents a significant confound on the other. It is conceivable that the national differences commented upon by brass players are simply the result of learning to play in a certain style; however, it could also be the case that different national schools or styles exist precisely because of playing differences that arise from being native speakers of different languages. The truth is likely to be found somewhere in the middle, especially given the fact that the concept also applies to other instruments where the influence of language is not readily conceivable, for example string playing.

To provide the reader with a better understanding of national schools or styles of playing, I am presenting a necessarily subjective account of the situation regarding brass playing in this section, which draws on my personal experience as a brass student learning to play the trombone in Germany. The subjectivity of this presentation is also necessitated by the scarcity of resources available on the topic (but see Antão & Moreira, 2015 for an intercultural approach to some of these issues).[17] I would like

[17] While I have tried, unsuccessfully, to find this kind of information online, I am sure that a considerable number of websites, especially blogs, contain isolated accounts of opinions on this matter.
The following statement by Argentinian conductor Daniel Barenboim, from an edited account of conversations with Palestinian intellectual Edward Said, provides an example of national differences in string playing: "… the fact that there is German art, a German style of playing, can be absolutely proven. There is a definite German string sound: a "German triplet" is a triplet where the first note of the triplet is drawn out to make a broad triplet; a "German upbeat" is a broad upbeat. Why? Because the language is like this. When you say *la lune*, the upbeat is short; when you say *der Mond*, it has to be longer and it has to be separated" (Barenboim & Said, 2004, p. 101).
That opinions on matters of national playing styles can be quite polarizing is evidenced by a recent "oboe hate mob" (Lebrecht, 2016) caused by an unfavorable review of a CD recorded by an American oboe player (Leung n. d.), in which the critic allegedly not only insulted the artist, but "an entire nation

to commence this section with a quote from London Symphony Orchestra Principal Bass Trombone Player, Paul Milner; I transcribed this information from a Masterclass video he recorded for the 2011 round of auditions for the YouTube Symphony Orchestra, which was open to players worldwide.

> I play for a British orchestra. … If you were to listen to a bass trombonist playing the excerpts from a Russian orchestra, French orchestra […] American orchestra, German orchestra, you name it, whichever different orchestra that is; we all have our own different international styles of playing. So, you might not think that the British way of playing is right but you know that, that's the way it goes (Milner 2010, mins 1:57-2:21)[18].

It is clear from Milner's statement that he believes in the existence of different (inter-) national styles of playing, and that playing in a different style might not sound 'right' to players coming from other backgrounds. I would certainly agree with Milner that, concerning trombone playing, there are differences in the playing by Russian, French, American, German, and British musicians.

Concerning French horn playing, Hageman's Doctor of Muscial Arts (D.M.A.) dissertation (2005) provides a similar list of national styles of playing; he does not mention a Russian style but rather assumes the existence of a distinct Viennese style, a very reasonable assumption given the use of a peculiar type of French horn design in Vienna[19]. Even though they might not translate completely to trombone playing, Hageman's descriptions of different national styles mostly agree with my impressions regarding trombone playing (see table 2.2 below). He outlines the overall situation regarding horn playing in his abstract:

---

or perhaps continent of oboe players as well" (Needleman, 2016). While the reviewer (who originates from, and lives in, Hong Kong) made it clear that he prefers the European style of playing (instruments also differ, causing characteristic differences in timbre), the following line seems to have been the trigger for Needleman's (over-)reaction: "Get this album if you want David Ludwig's Pleiades or if you are a fan of the American style of oboe playing" (Leung, n. d.).

[18] I tried to contact Paul Milner via Facebook (I could not find his email address), asking him about other sources that might provide information similar to his statement; unfortunately, he did not reply to my message.

[19] The Viennese horn is essentially a natural horn "with a valve section" built of rather simple 'Vienna valves' or "Wienerpumpen," which were used "all around Europe until about 1830," but today are found only on Viennese Horns (Hageman, 2005, p. 62; cf. Pizka, 1986). It also features a "larger bell than even some German horns," giving it a "warm quality" and a "mellow sound"; the negative side of this "most efficient" acoustical design, however, is that "[t]he valve-action is extremely slow, making technical, rapid passages extremely difficult to play" (Hageman, 2005, p. 65).

> Today, there are specific national styles of playing that have molded themselves into a form of national pride as to how the instrument should be played. In the last half-century, these styles have regionalized themselves, taking on vast differences even within national borders, none more so prominent than in the United States. […] Until the 1950s, national styles were very distinct and one could easily identify the differences between the different schools; however, the world has homogenized. Players everywhere are expected to be as good and sound the same, if necessary, as a horn player on the other side of the world (Hageman, 2005, p. ix).

I agree with Hageman's assessment regarding homogenization; while he does not specify whether this process is happening in a certain direction, I perceive a trend towards the American style for the trombone, or low brass playing in general (including tuba and other instruments less commonly played in the orchestra such as the euphonium; cf. Weber, 1978). However, this movement may also have slowed down a bit in the last ten or fifteen years, at least in Germany.

Hageman's thesis contains sufficient detail on the individual national styles, some of which I am reproducing here. He describes the French tone as "light with a distinctive vibrato," while the German tone is "thick, dark, and resonant with no vibrato," and the English tone is "purer, with no vibrato", further explaining that these differences in European horn playing "were initially the result of the type of instrument being played, which had been the case even in the natural-horn era" when larger instruments were played in Germany, compared to the instruments used by French players, "based on the tradition of the early Bohemian influence" (p. 7). Later on he states that "[t]he style of the English school of playing is literally a mixture of the German and French styles"; while "[t]he historical tone of the English players" would lean more towards the "French school," due to using French style instruments, it had always "been darker than the French, even when both schools were practically playing the same instrument" (p. 51). In the final chapter of his dissertation, Hageman describes the American school of French horn playing as having developed out of a "diverse conglomerate of individuals with backgrounds from all of the European schools of playing, each of which had a strong impact on the location in which they settled" (p. 66). For this reason, a "general American horn sound" applying across the entire country would not exist (p. 66); more recently, the "American sound" had become "more heterogeneous than ever before," often related to the different brands of instruments being played in various major orchestras (p. 73).

During my time as a bass trombone performance major at the University of Music and Performing Arts in Mannheim and at the University of Music Saar in Saarbrücken, the concept of playing in the German style often came up. Advanced students relating their experiences at professional auditions frequently talked about job applications in the former German Democratic Republic requiring players to perform on a German instrument, or if that was not an explicit requirement, players using American-style trombones dropping out of the competition as soon as a screen ensuring anonymity during the early rounds was removed (cf. Weber, 1978)[20]. Traditional German trombones feature a more conical bore than American instruments and a flat ring of added metal at the bell ("Kranz" in German); this leads to a sound that is not as rich in harmonics as the sound produced by American trombones (Weber, 1978, p. 569). Ironically, the modern American trombone itself was actually developed out of a truly innovative design developed by German manufacturers at the beginning of the 19th century to address the insufficiencies (small bore, limiting sonority and intensity) of an instrument that had remained virtually unchanged since the Renaissance (Weber, 1978, pp. 566, 568). However, when the Americans started copying this German design at the beginning of the twentieth century, they went even further than the specifications they were replicating, eventually ousting the design that had initiated the movement towards bigger instruments. This new 'American trombone' also finally managed to supplant the small-bore instruments remaining in use in France and Britain until the 1950s (Weber, 1978, pp. 567-568). This process of internationalization and convergence towards the American style of brass playing received another boost with the success of the Chicago Symphony Orchestra's (CSO) famed brass section under conductor Sir Georg Solti (1969-1991; Rosenthal Archives of the Chicago Symphony Orchestra, n. d.). Following in the footsteps of the CSO brass section were other changes such as a trend towards playing instruments with ever-larger bores and mouthpieces that increased the maximum dynamics and performance at loud volumes. Critics say that this has led to a change in playing character within the low brass section, with American tenor trombonists sounding more like bass trombonists,

[20] That this issue is conflated with personal opinion becomes clear from anecdotes about similar treatment of a bass trombonist playing an instrument with a silver-plated bell (allegedly leading to a brighter timbre).

and bass trombonists sounding like the tuba players of earlier generations (cf. Weber, 1978, pp. 569-570). Within the last ten years there seems to have been a bit of retraction from the most drastic of these advances, at least in Germany.

As the comparison of my individual impressions with the findings of Hageman's thesis shows, each brass instrument might have a somewhat different story regarding national styles of playing, which can be related to influential performers and/or instrument design. For trumpets, one can distinguish rotary and piston instruments, named after the style of valves they use. Rotary trumpets are used in the German-speaking countries, with Viennese instruments differing in bore diameter; instruments with piston valves are used in the USA and the UK, and for jazz and wind band playing in Germany. There is also a special instrument, often required in French compositions, called cornet-à-piston; this instrument is also used in British-style brass bands. Furthermore, there are German (rotary) and American (piston) versions of the Flügelhorn (spelled Flugelhorn in English). French horns, though not differing in terms of valves (apart from the Viennese tradition; cf. footnote 19), may differ in terms of bore, taper and bell shape (cf. Hagemann, 2005). Regarding the tuba, there are varying preferences as to whether differently-sized instruments are used for higher and lower parts; in Germany, players mostly use instruments in F and Bb, corresponding to Eb and Bb basses in the UK, while in the USA and Spain players prefer to use instruments in C exclusively; German instruments usually have rotary valves, while British instruments use piston valves almost exclusively, and C tubas can have either or a combination.

The following table provides an overview of various stereotypical features associated with the national styles mentioned in this section; this information is based on my own perception and conversations with fellow brass players of various backgrounds, and of course the dissertation by Hageman (2005). Note that I have left cells open where I am lacking the respective information.

	instruments used			
national style	**trombones**	**other brass instruments**	**sound/timbre**	**articulation**
German	mostly American, some traditional German; also American-style built by German manufacturers	rotary trumpets; piston trumpets for wind bands, jazz	full, dark, but not as extreme as American style	harsh', 'heavy'; full note-lengths
American	American (various manufacturers)	almost exclusively piston trumpets	dark, fluffy, 'big sound'	soft, imprecise, notes not very well separated, percept of legato
British	American-style instruments by British manufacturers, American	almost exclusively piston trumpets	brassy at loud volumes	strong; short note-length
Australasian (Australia, New Zealand)	American-style; smaller instruments sometimes used for pre-Romantic literature	piston trumpets for American and British literature; rotary trumpets often used for German literature	more similar to British than American style	more similar to British than American style
French	small-bore American-style instruments by French manufacturers	?	thin, bright; (historically extensive use of vibrato, now discontinued; cf. Barboteu 1975/2000)	fast, precise
Viennese (Austrian)	?	Viennese horn, Viennese trumpet (?)	?	?
Russian	?	?	loud, edgy	?
Spanish, South-American (Romance languages)	?	?	energetic playing	fast, precise

Table 2.2: National styles of brass playing.

These styles have often been associated with national traditions which in turn often fall together with physical borders and the languages spoken in these areas. It is, however, also possible that differences could co-occur with dialect distinctions, as exemplified by the perceived differences between American and British players. Most importantly, the association of regional distinctions with the native language and/or dialect of their speakers is somewhat arbitrary; differences could simply have arisen out of traditions that are imparted through the various (national) schools of playing. This thesis aims to evaluate empirical evidence to determine whether language differences might at least be partially responsible for the differences described above.

2.5 Previous empirical research on brass playing[21]

Given the long history of the use of speech syllables in brass teaching, it is likely that every brass teacher has at some stage wished they could directly observe what goes on inside a student's mouth and throat. Unfortunately, the articulatory and/or sound-producing movements involved in brass playing are hidden from view inside the oral and pharyngeal cavities of the player. A number of other visualization techniques exist, however, that allow the observation of the physical behaviors involved in brass playing. This section will provide an overview of these techniques and their application to brass playing since 1954, offering a comprehensive review of what is currently known about the vocal tract movements happening during brass playing, and providing information about the articulators that could contribute to language influence on brass playing.

2.5.1 X-ray studies 1954-1975

The earliest technique that could answer the desire to document the physical changes involved in brass playing was x-ray imaging, or radiography, as it was called in its early days. It was applied soon after that became feasible and resulted in a number of dissertations that were completed using various developmental stages of the technique (still frames to moving images). Unfortunately, the discovery of the risks involved with x-ray exposure soon rendered further research too dangerous and the demise of the technique also brought with it the end to a promising string of early empirical research. The early findings discussed in this section were, if not forgotten, at least never seriously considered by the brass playing community (cf. Irvine, 2003). Hall's Ph.D. dissertation (1954), though limited to the capture of still frames, stands out from all other work discussed in this section due to the fact that Hall was the only researcher to control for variation introduced by different instruments, as well as

[21] Most of the material covered in this section has previously been published in Heyne, M., & Derrick, D. (2016a). Visualization techniques for empirical brass instrument research. *Journal of the International Trumpet Guild, 40*, 6-14, 24.

making objective measurements of the produced sound quality with an early ‘sonograph’; this technique had previously been applied by Gibson in his MA thesis (1942), which analyzed acoustic data only. All of Hall’s participants played selected tones on a ‘control trumpet,’ a “Rafael Mendez model” with a “Mendez mouthpiece” manufactured by a “well-known’ manufacturer” and not used habitually by any of the participants (Hall, 1954, p. 267), in addition to their own instruments. ‘Sonograms’ of the notes produced in all conditions were reprinted in the thesis; unfortunately, the scanned quality of the digitized document renders the extraction of his sonographic data impossible and restricts the readability of his x-ray findings to schematic tracings reproduced at small size. The main findings of Hall’s study were that different participants used unique individual positions of the tongue and jaw during playing, and that individuals tended to be consistent in using “the same basic formation in every register,” indicating that modifications while changing registers “were not large”: “[i]n nearly all cases, the changes were not as great as the changes between extreme vowels” (pp. 246-247). All of Hall’s tongue tracings were taken in the midsagittal plane and images taken during the spoken production of the extreme vowels[22] “(ah)” (/ɑ/), “(oo)” (/u/), and “(ee)” (/i/) in three different pitch ranges (p. 27) allowed him to compare these tongue positions to the ones utilized while playing. The most “common formation” used during playing “was that of ‘a’ (ah)” but the author added that “other players used the ‘u’ (oo) formation or intermediate formations between these extreme vowels” (pp. 246-247).

Subsequent studies by Meidt (1967), Haynie (1969; cf. Haynie & Hardin, 2007), Amstutz (1970; cf. Amstutz, 1977), Frohrip (1972) and De Young (1975) largely confirmed Hall’s findings, in addition to observing a wider range of playing conditions that included changes in dynamics and tongue placement for different types of articulations. Participants for some of these studies included players of various brass instruments, while Frohrip (1972) and De Young (1975) observed trombone players exclusively. With all the individual variation observed in these studies it is not surprising that Meidt (1967) reported a difference in results compared to Hall’s (1954)

[22] Note that Budde (2011) incorrectly states that the vowels produced by Hall’s participants were sung – instead Hall’s dissertation reads: “with the instrument at his lips” each “subject said the following extreme vowels, prolonging each sound for three seconds to permit obtaining of the data (Hall, 1954, p. 27). Also note that “(ah)” and “(oo)” were incorrectly transcribed as “A” and “O” in Zsaisits’ MA thesis (2002) but correctly transcribed as /ɑ/ as in American English *pod* and /u/ as in *booed* by Budde (2011, p. 81).

findings with respect to register changes: among his participants, some players displayed large changes in tongue position with “the variations in formation […] usually approaching, if not actually reaching, the extreme “ah” and “ee” vowel formations” (p. 66).

One other early study that deserves specific mention is Hiigel’s dissertation (1967) on *The relationship of syllables to pitch and tonguing in brass instrument playing*. This researcher asked his participants to ‘think’ prescribed syllables printed underneath the music while performing selected notes and found no evidence “that thinking a syllable during performance will tend to simulate the tongue position resulting from the enunciation of that syllable” (p. 108). Similarly, significant differences were found “between the tongue placement for performance of the various pitches and styles and placement for the enunciation of the syllables” recorded separately, even for the players who claimed to use those specific syllables during playing. The overall tendency was for the “tongue arch” to be placed higher with the tongue tip “farther forward” when compared to recitation (p. 107).

X-ray imaging was also used by Carter (1966/69) to observe the role of the larynx during brass playing. He found that during brass playing “the size of the glottis opening varies with loudness level, being small for soft playing and large for loud playing” and that “there is no practical difference in glottis opening from the high to the low register, indicating that pitch control *within the normal playing register, can be discarded as a possible function of the variable glottis aperture*” (p. 427; italics in the original). Carter’s findings were partially supported by a study for the U.S. Department of Health, Education, and Welfare (Nichols et al., 1971) initiated by Fay Hanson, which presents an early attempt at simultaneous measurement of various parameters involved in trumpet playing. These authors found the “control movements of the true vocal cord” to be “the most important mechanisms for the production of sound interruption in trumpet playing” with an opening “from 2-3 mm whenever a sound was produced”; this applies to the intermediate and advanced players in their sample, a beginning player displayed “deficient and unpredictable laryngeal movement” (p. 37). While these studies suggest conscious control of the vocal folds to regulate airflow (Carter made explicit reference to a theory put forward by Farkas, 1956), later studies employing laryngoscopy have shown that changes in vocal fold position are rather small and moreover “self-adjusting or involuntary” (Bailey, 1989, p. 105; cf. sources listed in section 2.3.2).

Finally, advances in x-ray technology have reduced the amount of radiation exposure so that modern technologies such as Cone Beam Computed Tomography (CBCT) can again be used to measure features of the vocal tract during sustained postures. Cilingir (2012) applied this technique to measure teeth malocclusion as well as oral and nasal cavity size in order to investigate their effect on trumpet performance proficiency. Results reported in Kula et al. (2015) indicate that although not all types of malocclusions necessarily influenced trumpet performance, specific factors such incisor inclination could negatively affect certain areas of skill, in this case articulation (p. 5). Overall, they conclude “that total compensation [for malocclusions] probably does not occur and that dental anatomic factors influence the quality of trumpet playing after years of practice” (p. 6).

2.5.2 Observations of the brass player’s lips

The vibrating behavior of brass players’ lips was first investigated by Martin (1942) in his seminal study using a specifically designed L-shaped mouthpiece and stroboscopic photography with a resolution of 50 frames-per-second (fps) while observing cornet playing. Weast (1963) expanded on Martin’s findings by including players of all kinds of brass instruments, using an improved experiment design with transparent plastic mouthpieces and a stroboscope disc to observe the different phases of lip vibration. He found “overwhelming evidence of the upper lip being the primary vibrating mechanism” (p. 223), while “the activity of the lower lip was highly erratic” (p. 222). Later studies (Leno, 1970; Copley & Strong, 1996; Bromage, Campbell, & Gilbert, 2010) have shown the possibility of the reverse pattern occurring for upstream type players (cf. Reinhardt, 1973), and improved frame rates and image quality have enabled a more detailed description of brass players’ lips. Most recently, researchers have used automated measurements of lip opening area and pressure within the mouthpiece throat and at the instrument’s bell to investigate the different mechanisms used to effect lip-slurring on brass instruments (Logie et al., 2010; Hruška, Švejda, & Guštar, 2015). Stroboscopic and kymographic (Švec & Schutte, 1996) methods can also be employed to diagnose playing problems such as embouchure dystonia (for a definition and an overview on embouchure dystonia see Frucht, 2009).

Lip vibration has also been studied using indirect methods of observation such as using a water nanometer to measure steady pressure in the player’s mouth compared with airflow velocity in the throat of the mouthpiece (Elliott & Bowsher, 1982), or employing a strain gauge (described as “force stripes” in Rosset i Llobet, 2005) to record lip movement in combination with a probe microphone to estimate sound pressure within the mouthpiece (Yoshikawa, 1995). Recently, researchers have also adapted the principle of electroglottography to brass playing (cf. papers discussed in section 2.3.1.2 above); this technology referred to as a “electrolabiograph” (Hézard et al., 2014, p. 6) was applied in Fréour’s Ph.D. thesis (2013) and a paper by Boutin et al. (2013). By placing two contact electrodes made of tin-plated copper foil on the upper and lower rim of a plastic mouthpiece and sending a high frequency modulated current through the lips, the contact quotient of the lips can be calculated in a way that is far less intrusive than the direct observation methods used in early studies.

2.5.3 Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) is the only modern visualization technique that surpasses the possibilities of x-ray imaging, although it does not image bone and teeth, and has some other more serious limitations. MRI scanners use very strong magnetic fields and radio waves to compute images of the human body so that no ferromagnetic materials can be taken inside the machine. This means that no real brass instruments can be played inside an MRI scanner, although this would also be problematic due to the small bore of the tunnel which subjects have to enter to ensure that the body part to be imaged is at the absolute center of the scanner. Another drawback of the technique is the fact that the most powerful machines with the best resolution require subjects to be situated in supine position which introduces small artifacts due to gravitational effects compared to the upright position normally assumed during brass playing (Tiede et al., 1997; Shiller, Ostry, & Gribble, 1999; Wrench, Cleland, & Scobbie, 2011; Traser et al., 2013; negligible according to Buchaillard, Perrier, & Payan, 2009). Additionally, loud repetitive noises inside the scanner arise from the need to rapidly switch on and off its magnetic coils, which could influence vocal and instrumental performance[23].

While the potential exists to use functional MRI (fMRI) to look at brain activation specific to brass playing or shared with other forms of human behavior, such as speech

[23] The so-called Lombard effect or Lombard reflex has been demonstrated to affect the vocal production in noise not only for humans but also for nonhuman animals (Lombard, 1911).

production, the most obvious application of MRI to brass research is real-time MRI, which allows the observation of the body in motion. Recent advances in the way MRI composite images are computed has increased the temporal resolution of the technique such that it can produce an image every twenty to thirty milliseconds with a spatial resolution of one-and-a-half to two millimeters relative to the scanned plane.

Kaburagi et al. (2011) used MRI to investigate the effect of a single player's vocal tract on trumpet sound, requiring the player to maintain the same vocal tract configuration for at least thirty seconds. The instrument used was a plastic replica built according to the specifications of a Yamaha student model (without any valves pressed down) with an acrylic mouthpiece. In addition to detailed measurements of the vocal tract for three different pitches, the paper also includes images of the vocal tract while producing three different Japanese vowels. The comparison of tongue positions used during playing with those of the sustained vowels indicated that "the tongue posture for the low and mid pitches was similar to that for the back vowel /o/", while "the tongue posture for the high-pitch trumpet sound was similar to that for the vowel /u/[24], but it was located slightly posteriorly" (p. 538). Wiggins and Storey (2010) used high frame rate FLASH (fast low-angle shot) acquisition to observe instrument-specific techniques such as circular breathing, interdental tonguing and "modifications of the vocal tract which cause a formantlike [sic] filtering effect in the sound" of a single didgeridoo player; the test instrument was made of fiberglass tubing with "a section of flexible corrugated plastic tubing" bent 90 degrees to enable use inside an MRI scanner (p. 5007).

A larger number of subjects playing different instruments were examined at the Freiburger Institut für Musikermedizin in Germany using real-time MRI. They produced a wonderful non-technical DVD-ROM with the title *Physiological Insights for Players of Wind Instruments* that features MRI videos explaining the tonguing and breathing movements involved in wind playing (Spahn et al., 2013). A more detailed study by Schumacher et al. (2013) conducted at the same institution observed the motor functions in trumpet playing. The main findings of this research were that: "1. With increasing tone pitch in octave jumps and in playing natural tones, there was an increase in total free space of both the oral and pharyngeal cavity. The increase of

[24] Note that the vowel /u/ in Japanese (more correctly transcribed as /ʉ/; Nogita, Yamane & Byrd, 2013) is thought to be more central than the comparable cardinal vowel.

both to achieve the higher pitch was greater in the pharynx than in the oral cavity." And "2. The increase in areas of oral cavity and pharynx are present also when switching from lower to higher loudness and when performing crescendo to decrescendo... However, no general difference in change of oral and pharyngeal cavity can be observed" (p. 1177). In both studies, simple approximations of brass instruments consisting of a non-metal mouthpiece attached to a common rubber tube with a sound funnel were used (p. 1172).

The most recent research employing MRI for brass research is a collaboration by American kinesiologist and French horn player Peter Iltis with researchers at the Max-Planck-Institut in Göttingen, Germany. This research was initiated by Eckart Altenmüller (head of the department of Music-Physiology and Musician's Medicine at the University for Music and Theater, Hannover), a renowned specialist on focal task-specific dystonias in musicians. After considering MRI data of French horn players with embouchure dystonia (cf. Frucht, 2009), it became clear that the pathology of this condition could only be determined by comparing it with similar data from healthy subjects. The current corpus of recordings includes MRI video of twelve (healthy) elite performers (recruited from the Berlin Philharmonic and other renowned orchestras) and five dystonic horn players (all "former professional musicians"; Iltis et al., 2016, p. 70). To render playing conditions inside the MRI scanner more realistic, all subjects used a custom-built, MRI-compatible horn consisting of a non-ferromagnetic bell with graduated plastic tubing covering the distance from just outside the scanner to the player's mouth (Iltis et al., 2015b, p. 375). A first paper published in 2015 reports findings comparing data for six elite players with that of five players with embouchure dystonia. While no differences between groups were found regarding "the resting position of the tongue, teeth, jaw and oral cavity" (Iltis et al., 2015a, p. 5), "significant differences in movement strategies" were found, which, according to the authors, "may provide insight into possible triggers for or consequences of embouchure dystonia" (p. 8). The "fairly consistent motor strategy" observed by the elite players "with respect to the anterior-dorsal aspect of the tongue when performing a slurred, ascending 11-note harmonic series" involved the following physical changes (p. 6):

> On the lowest notes, the tongue was positioned low within the oral cavity creating a large cavitation until the 5th harmonic was reached. Subsequently higher notes involved a progressive upward and forward movement of the dorsal surface of the tongue that decreased the size of the oral cavity (p. 6).

The conformity among elite players regarding this strategy is remarkable given earlier accounts of widespread individual variability[25], and clearly visible in a pedagogical video featuring data by 10 elite players, available on YouTube (Epstein & Iltis, 2016b). In both the 2015a paper and the YouTube video, the strategy is described as producing “adjustments […] similar to those used when phonating the English vowel sounds” (Iltis et al., 2015a, p. 6), and is then used to justify the “purposeful use of these vowel/tongue adjustments by several horn teachers” (p. 6). While there is similarity to a varying extent among the various elite players concerning tongue shape in the oral cavity (the authors’ analysis considers midsagittal measurements only), pharyngeal constriction during horn playing (as displayed for three elite players in the YouTube video by Epstein and Iltis, 2016b) differs markedly from the recommended vowel tongue positions[26]. Notwithstanding, employing vowel tongue shapes in teaching could lead students to replicate successful patterns such as narrowing the airway which “results in acceleration of the air column and higher vibration frequencies of the lips” (Iltis et al. 2015a, p. 6; citing a consideration by Epstein).

Returning to the findings for players suffering from embouchure dystonia, the authors report that they “display smaller oral cavities on the lower notes, and a less precipitous reduction in cavitation as they move to the higher notes,” although they were also overall “less consistent across subjects” (Iltis et al., 2015a, p. 6). Similar findings are reported in the 2016 paper featuring increased participant numbers (twelve elite and

[25] Hiigel’s study (1967) is the only other study using professional subjects exclusively, and the author seems to have pursued a similar aim by representing “the ultimate level of proficiency in the art of brass instrument performance” with the expectation that these players would “provide a model of the correctly functioning oral physiological structure during performance” (Hiigel, 1967, p. 22). Subject numbers for this study, however, were low (two each for trumpet, French horn, and trombone) and although the dissertation does not contain individual comparisons of players performing on the same instrument, the summary reports “significant differences in tongue placement” existing “between subjects in all measurement dimensions” (Hiigel, 1967, p. 19). Note also that the author makes the following recommendations to “(1) avoid the use of syllables as simulators of correct tongue placement for brass performance; (2) teach for a more consistent tongue placement, both tip and posterior arch, for the total pitch range of the brass instruments…” (Hiigel, 1967, p. 110). Of course, roughly 50 years lie between these two studies and it could be that brass playing overall has improved significantly, leading to a further refinement of top-level performance technique.

[26] Note that the authors later on qualify their claim in relation to Epstein’s pedagogical concept: “Though the degree to which the elite performers adhere to such an approach is not as systematic throughout the range as Epstein recommends, it is apparent that the general pattern is present, particularly in the upper range of the instrument” (Iltis et al., 2015a, p. 6).

Note also that I could not find any information regarding the elicitation of vowel tongue shapes in the papers or YouTube video. I assume, however, that vowels were produced in isolation and sustained. Findings from articulatory phonetics show the variability of tongue shapes (and acoustic data) across tokens produced in various phonetic contexts, questioning the appropriateness of using single measurements in comparisons such as these.

five dystonic players) and additional measurements of jaw displacement and mean cross-sectional area (XSA). In terms of jaw movements, dystonic players showed a pattern of raising and then lowering the jaw for an ascending slurred eleven note harmonic sequence (the position used for the lowest note served as reference for all measurements), while the elite players displayed a "progressive elevation of the jaw […] with little change" occurring during the final five notes of the sequence; in the discussion, however, the authors add that "[c]are must be taken in considering the apparent jaw lowering of the last harmonics […] as there was considerable variability in the ED [embouchure dystonia] response" (Iltis et al., 2016, p. 74). For the same sequence of notes played in descending order no divergent pattern was found, although the elite players showed a "a trend toward a greater change" throughout the sequence (p. 73). The mean cross-sectional area at the tongue dorsum was calculated using coronal RT-MRI (real-time) films by "choosing a frame that occurred when the sound of the note was stable, and when the size of the ROI [region of interest = "air channel formed between the dorsal surface of the tongue and the roof of the mouth" (p. 74)] appeared to be stable and consistent for several seconds. This usually occurred at approximately the midpoint of these exercises" consisting of long notes (high Bb4 and low Eb2) sustained at pp and ff dynamics (p. 72). No information is given in the paper relating to the exact orientation of the coronal image slices used for these calculations; images of two players included in the paper, however, suggest the orientation to be roughly perpendicular to the players' occlusal planes. Here, clear differences are apparent between a reduced number of elite and dystonic performers (four each); when

> "transitioning from high to low notes at the same dynamic level […], the elite performers tend[ed] to increase the size of the air channel formed between the dorsal surface of the tongue and the roof of the mouth on the high notes, while the opposite is true for the subjects with ED" (pp. 73-74).

The authors describe further discrepancies for changes occurring with an increase in intensity (pp versus ff). Incorporating the three quantified parameters, the authors interpret the behavior displayed by the dystonic players as an inefficiency contributing to embouchure dystonia; nonetheless, they stress that this leaves the question unanswered whether "these aberrant lingual movements" are "causal triggers for ED or simply manifestations of broader-based maladaptive sensorimotor processing" (p. 76).

The same core group of authors published two other papers in 2015, one assessing "the efficacy of RT-MRI films" at different frame rates to measure "double tonguing performance" (Iltis et al., 2015b, abstract), while another one investigated the oral cavity changes in performers of different brass instruments (Iltis et al., 2015c). The paper on fast tongue articulations showed that "acquisition rates of 30 fps are inadequate to accurately detect tongue movements during double tonguing, but that rates of 100 fps do allow for a precise quantification of movement," demonstrating for the first time "the extreme performance of elite horn players" (Iltis et al., 2015b, abstract). "A regular sinusoidal anti-phase pattern for the tongue tip and midtongue is readily observed" with the faster acquisition technique (corresponding to what the authors label "T" and "K" articulations; Iltis et al., 2015b, p. 379), which can be rendered more descriptive by applying a Fourier transform. Finally, in the study observing players of different brass instruments, one subject each for trumpet, trombone, French horn, and tuba were recorded performing an ascending pitch sequence over two octaves on a B.E.R.P. practice device (Musical Enterprises, Fairfax, California) that is meant to simulate the aerodynamic resistance afforded by an actual brass instrument. The real-time image acquisition rate was thirty Hertz, which allowed the coarse analysis of double-tonguing, although no differences between the different instruments could be observed. In terms of the overall tongue movement patterns, however, an opposing pattern emerged for trumpet versus trombone and, to a smaller extent, for French horn versus tuba players: "For the trumpet, the anterior area change … was characterized by a forward and upward projection of the tongue at both extremes of the register while for the trombone, a marked forward and upward shift of the tongue occurred when moving between the lowest … to [sic] the highest … notes" (p. 140).

2.5.4 Ultrasound imaging of the tongue[27]

The first application of ultrasound tongue imaging (UTI) to wind playing known to the author was by Wein and colleagues (1989) at the Radiological Hospital in Prague (then Czechoslovakia), with data analysis carried out in Aachen, Germany. A summary of their research on wind playing published in 2000 (Šram & Svec, 2000)[28] reports the following regarding the use of the tongue during trumpet and tuba playing:

[27] See chapter 6 for a technical description.
[28] It is not clear from the text whether these findings are based exclusively on ultrasound research as the authors also used x-ray imaging before switching to ultrasound. The publication also discusses

> The configuration of the speech organs for a tone played in middle and lower register approaches a position similar to that of the organs of speech when pronouncing the vowel “A” … The position of the articulators when playing in the high register approaches, however, the position of the vowel articulation “Y” or “JÍ” (p. 155; translation by myself)[29].

Later on in the text they elaborate on this, stating that the articulation of sounds – vowels, consonants and syllables – is only approximated, and has individual variations on the interaction of the individual instruments and individual musicians (p. 155). Another important function of the tongue mentioned in this publication but not reported in any other study (possibly because it might be too obvious) is its use in helping to continuously moisten the lips and thus enabling their regular vibration. Finally, these authors also describe the vibrating tip of the tongue as a significant factor for sound production on brass instruments, especially on the tuba, and the usage of rhythmic movements of the tongue base and the walls of the larynx to affect vibrato (p. 156).

Another application of UTI to brass playing was in a study by Zielke (2010) at the university hospital in Düsseldorf, Germany. Tongue and motor activity of the neck and face were observed for players of different wind instruments, with the intent of quantifying and describing these behaviors for the purposes of music medicine. The author found that ‘tongue amplitudes’ (meaning the displacement of the tongue during certain movements) were larger for players of brass instrument than flutists, for loud versus soft playing and for attacked notes versus slurred articulation. While we could find no information on how Wein et al. (see above) kept the ultrasound transducer in place underneath the participants’ chins, Zielke reports that one of the investigators held the transducer under the chin for her participants while two other investigators performed the ultrasound and video recordings (p. 21). Although this might be sufficient for measurements of displacement, fixing the ultrasound transducer in place under the participant’s chin becomes a necessity if different tongue positions are to be compared, such as in this thesis. This issue will be further discussed in the ultrasound methodology section.

findings gained by using videostroboscopy and videokymographie to observe the brass players’ lips in motion.

[29] I believe that the pronunciations indicated by the use of Czech letters refer to the following vowel positions on the IPA chart: “A” = /a/, “Y” or “JÍ” = /i~ɪ/ (cf. Šimáčková, Podlipský & Chládková, 2012).

Professor Angerstein, supervisor of Zielke's dissertation, is a phoniatrician and audiologist who runs a weekly consultation hour for wind players with tonguing issues that uses UTI as a diagnostic and therapeutic tool[30]. Although not reported in any scientific publications, an article describing his work in German wind players' magazine *Clarino* reports an important observation gained from the coronal application of lingual ultrasound: normally, the tongue displays a groove along its midline, whereby the air is channeled from the larynx to the front teeth. If this channel trails off to the side, the air misses the orifice of the instrument, which then leads to problems of articulation (Härtel, 2012, p. 22; quoting Professor Angerstein; translation by myself). Angerstein, himself a didgeridoo and alphorn player, seems to be convinced of vocal tract influence on brass instrument sound: the tongue muscles control, together with the respiratory muscles, the larynx, the cheeks and the soft palate, the airflow during wind instrument playing and thus are very influential, among other things, in determining sound quality. Furthermore, tooth position, palate shape, size and position of the tongue play an important role regarding sound quality (Härtel, 2012, p. 22).

2.5.5 Electromagnetic articulography

Electromagnetic articulography (EMA) uses alternating current magnetic fields and electromagnetic detection to track the positions of small metal transceiver coils attached to the articulators. Modern equipment can track the position of the sensors in three-dimensional space and can factor out head movement by attaching some of sensors to anatomical landmarks such as the mastoid processes. EMA offers excellent spatial (amount of detail) and temporal resolution (number of measurements per second) but can be sensitive to changes in temperature, air movement, and local magnetic fields. Additionally, applying the sensors is time consuming and invasive, and they may come loose during data collection.

The only documented application of this technique to brass playing so far is Bertsch and Hoole (2014); this was a pilot study with a single participant to demonstrate the usefulness of the technology for brass research[31]. Interestingly, playing a regular brass trumpet did not seem to impair the magnetic field required for the data collection;

[30] For more information see http://www.uniklinik-duesseldorf.de/en/unternehmen/kliniken/phoniatrics-and-paediatric-audiology/

[31] The system used in this study was a Carstens AG501 by Carstens Medizinelektronik, Bovenden, Germany. Video from the data collection is available at https://www.youtube.com/watch?v=bAlYUt3FGaQ&feature=youtu.be.

the mouthpiece used by the participant was a plastic replica of his regular mouthpiece. As part of this study, I also recorded a trombone player using a different EMA system (our system at NZILBB is a Northern Digital Inc. Wave machine) and synchronized UTI. Although we asked our participant to play a plastic trombone with a plastic mouthpiece for the main part of the experiment, we asked him to briefly switch to a brass trombone with a brass mouthpiece at the end. Preliminary analysis suggests the brass had but a small impact on measurement accuracy, increasing the standard deviation of the residual recorded motion of the head motion tracking sensors by about 0.13 mm (from 0.43 mm to 0.56 mm) compared to that from the same piece played with the plastic trombone. Nevertheless, it must be noted that these sensors were located several centimeters away from the brass mouthpiece – metal may cause more variation in the tongue sensors as they were closer to the mouthpiece.

2.5.6 Conclusion

This review of previous empirical research on brass playing using various visualization techniques shows that researchers have tried and succeeded at measuring some of the physical changes occurring in performers, making use of the most sophisticated techniques available at the time. Perhaps the most striking finding is the lack of any patterns that apply reliably across players of different instruments, proficiencies or even individuals. Furthermore, it is unclear how different physical changes, e.g. at the glottis and lips, might interact; note though, that a few studies have tried to observe multiple factors at the same time, such as Nichols et al. (1971), Logie et al. (2010), and Fréour and Scavone (2011). Of the described visualization methods, MRI, UTI and articulography allow the observation of major vocal tract movements in real time, meaning that they represent viable options to address the research questions posed in this thesis. However, only UTI satisfies the requirements of this specific study, as it is currently not possible to fit a trombone into an MRI scanner, and the invasiveness and time demands of articulography prohibit the observation of a sufficient number of individuals.

Although the L1 of the participants was not reported in any of the studies reviewed above, the assumption would have to be that the majority were speakers of AE, given that most of the research was conducted in the USA. We would thus expect plenty of individual variation among players with a shared L1 even if there was no influence of L1 on brass playing. Such individual variation, however, is also common in speech production (cf. section 3.3.4 below), and comparing this behavior to vocal tract behavior during brass playing might show that there are common factors leading to individual variation affecting both domains; such a comparison will be presented in chapter 3. Before that, however, this chapter will conclude with a description of research on the perception of brass instrument sound, and a summary of the only two previous studies on language influence on brass playing.

2.6 Perception of (brass) instrument sound by expert and non-expert listeners

Several studies have shown that musical instruments are more easily distinguished by both expert and non-expert listeners when attack and release of notes are included with the test stimuli (Thayer, 1974; Elliott, 1975; Paul, 2005; among others). This effect seems to be related to the characteristic transient phenomena of the various groups of instruments; the term transient refers to the beginning (and end) of complex tones in which the spectrum of individual overtones changes rapidly, "as opposed to the relatively stable, steady state portion of the sound that follows" (Spiegelberg, 2002, p. 54). Transient phenomena have been shown to be particularly relevant as a characteristic cue for recognizing brass instrument sounds (Thayer, 1974). Winckel (1967) pointed out that

> [i]n the case of the trumpet, the 1st and 2nd partials develop rather quickly, while the upper partials require more time to develop their full energy; in the case of the violin, the first two partials develop more slowly than the upper partials. This is the reason why the trumpet sounds more clearly defined, with more fundamental, than the violin (p. 32; drawing on research by Backhaus, 1932).

Winckel takes his claim to be supported by a close resemblance between the onset characteristics of wind instruments and those of "certain speech phenomena" (Diaz, 2011, p. 30); his concrete example shows the onset amplitude curves of the first five partials of the "speech sound complex 'dah' which he recognizes as resembling more closely "the sound of a trumpet, in the sense of a plosive, than […] that of a violin" (p. 33)[32]. Recent findings by Paul (2005) suggest that although perceptional abilities differ

[32] Plots showing the onset of the first five partials of "dah" and various musical instruments are printed on the corresponding pages of the book.
A similar correspondence has also been noted by Patel & Iversen (2003) for North Indian tabla drumming. They tested the hypothesis that 'vocables' (nonsense syllables used to name drum sounds) represent a case of sound symbolism (onomatopoeia) and found that the acoustic properties of the

little between trained listeners at High School and graduate student levels, extraordinarily high levels of recognition (91 percent) were achieved in identifying one's own instrument's timbre (p. 24).

Advances in neuroscience have enabled scientists to investigate more closely the mechanisms involved in music perception by observing the human brain during, for example, the perception of complex tones. The so-called 'missing fundamental' phenomenon was first described by von Helmholtz (1877/1954), who distinguished a 'synthetical' mode of perceiving the fundamental pitch of complex sounds, as opposed to an 'analytical' mode based on the perception of single partial tones (p. 62). Modern studies (Preisler, 1993; Seither-Preisler et al., 2007) have shown that individuals tend to have a preference for either mode of perception, explained using different terminology by Schneider and Wengenroth (2009):

> Holistic' or 'synthetic' listeners recognize the sound as a whole, and appreciate its pitch and timbre as characteristic qualities of the entire sound [von Helmholtz's synthetical mode], [while] 'spectral' or 'analytical' listeners break up the sound into its harmonic constituents, at the expense of timbral qualities of the sound as a whole [von Helmholtz's analytical mode] (p. 315).

These studies, however, also indicate that there are "intermediate listeners [perceiving] holistic and spectral cues simultaneously to varying degrees" (Schneider & Wengenroth, 2009, p. 315), and that results can be influenced by "certain stimulus variables [with] predictable effects on responses" (Ladd et al., 2013, p. 1396)[33]. Neuroanatomical measurements have revealed that there exist "enormous inter-individual differences in terms of shape, gyration, size and number of duplications of HG [Heschl's gyrus]" (Schneider & Wengenroth, 2009, p. 319). Interestingly, non-musicians usually have a small, single Heschl's gyrus, while professional musicians frequently display a doubling of size and one or more reduplications; the lateral Heschl's gyri (also called transverse temporal gyri) are parts of the temporal lobes hosting most parts of the primary and secondary auditory cortex; these are important areas for music and sound processing (Schneider & Wengenroth, 2009, p. 319).

---

drum sounds in question were reflected by "a variety of phonetic components" of the vocables (p. 925). Furthermore, their perceptual experiment showed that naïve listeners could match the vocables to the corresponding tabla sounds.

[33] Pantev et al. (2003) demonstrated that participants can also be trained to perceive "virtual pitch" ('holistic' mode in von Helmholtz's terminology) even if they are preferentially spectral listeners and that such "relatively short-term training can lead to plastic changes in the primary auditory cortex" (p. 442).

Of particular interest to this thesis are the findings from a large-scale study by Schneider and colleagues (Schneider et al., 2005a&b) for which they tested 421[34] right-handed[35] subjects regarding their pitch perception preference, including 373 musicians; a subgroup of 87 subjects was investigated in more detail using MRI scans of brain structure and functional MEG (Magnetoencephalography) of neural activity in response to complex tones. A "strong correlation between pitch perception preference and asymmetry of brain structure and function in the pitch-sensitive lateral areas of Heschl's gyrus (HG)" was found, "irrespective of musical ability" (Schneider et al., 2005b, p. 387). Furthermore, the data show an association of pitch processing preference with instrument preference: "[F]undamental pitch listeners" exhibited larger "gray matter volume … and enhanced P50m activity … in the left lateral HG, which is sensitive to rapid temporal processing," and reported that they played "percussive or high-pitched instruments that produce short, sharp, or impulsive tones (e.g., drums, guitar, piano, trumpet, or flute)" (Schneider et al., 2005b, p. 387). In contrast, spectral pitch listeners exhibited a "dominant right lateral HG, which is known to be sensitive to slower temporal and spectral processing," and were players of "lower-pitched melodic instruments that produce rather sustained tones with characteristic changes in timbre (e.g., bassoon, saxophone, french horn, violoncello, or organ)" (Schneider et al., 2005b, p. 387). Although not mentioned explicitly, the trombone would belong to the latter group, suggesting that a majority of trombone players likely perceive pitch spectrally[36], making them highly sensitive to the spectral content of complex tones, and trombone notes in particular. This further suggests that trombone players, perhaps more so than trumpet players, have the perceptual ability to notice subtle differences in the timbre produced by their instruments, increasing their awareness of the possible impact of variations in vocal tract configuration on trombone sound. In turn, they might be more inclined to experiment (whether consciously or subconsciously) with various vocal tract configurations, possibly mitigating the hypothesized effect of native language on trombone playing.

[34] Participant numbers differ slightly across papers but it is obvious that at least some of the subjects were the same individuals. Thus, I am reporting the highest cumulative number from Schneider et al. (2005a).
[35] A weak association of pitch processing preference with handedness was reported in a paper by Laguitton et al. (1998).
[36] A short test ("Kurztest" in German) to determine one's pitch processing preference is available online at http://www.musicandbrain.de/kurztest.html (last accessed 14 July 2016). My own result fits the prediction: extreme spectral pitch listener ("Extremer Obertonhörer").

Note, however, that it is not yet clear whether pitch processing preferences are "shaped by intense training or rather reflect innate, genetically determined predisposition"; this "remains a matter of unresolved debate" (Schneider & Wengenroth, 2009, p. 315). In their early work, Schneider et al. (2005b) seemed to presuppose genetic factors: "It is likely that both magnitude and asymmetry of lateral HG, and the related perceptual mode, may have an impact on preference for particular musical instruments and on musical performance" (p. 387). Evidence for an association of processing preference and speaking a tone language, however, points in the opposite direction (Petitti & Perrachione, 2015; cf. Wong et al., 2008). Similarly, findings by Pantev et al. (2003) showing that "highly skilled musicians exhibit enhanced auditory cortical representations for musical timbres associated with their principal instrument, compared to timbres associated with instruments on which they have not been trained" (p. 4), also point to a substantial influence of experience rather than genetic predisposition. Of the two groups of musicians investigated in Pantev et al.'s study (2001 & 2003), one consisted of professional trumpet players. Their MEG data showed "a robust timbre-specific enhancement of auditory cortical representations" (2003, p. 5) for trumpet tones, which differed from their perception of violin tones (the reverse was true for violinists; 2001, p. 4)[37]; furthermore, "[o]verall, larger responses were recorded to the musical stimuli among the trumpeters" (2001, p. 5). The authors suggest that "[c]ross-modal feedback arising from" the use of the "pharynx, larynx, tongue, lip, and diaphragm to produce musical sounds" might be responsible for an "[a]ugmentation of the cortical representation for trumpet tones in trumpeters," compared to violin players, although an alternative explanation could be that "[t]rumpet tones are ... typically played more loudly than … string tones, which may magnify the representation for these tones and other timbres that are heard during musical performance" (2001, p. 5).

A limited number of studies have investigated the ability of expert (and non-expert) listeners to perceive small differences in brass instrument sound. Diaz (2011) asked 31 brass instrument and 32 other college music majors to rate pairs of trombone notes according to whether they perceived their articulation getting softer, harder, or staying the same. The stimuli were recorded by a professional trombonist instructed to

[37] This result may help explain why extraordinarily high levels of recognition (91 percent) could be achieved in identifying one's own instrument's timbre in Paul's study (2005, p. 24).

produce “dah,” “tah,” and “pah” attacks on the note F4, which was chosen “due to its relative ease of execution” (p. 31). Diaz’s results suggest that listeners can discriminate different articulations at statistical significance even when the “amplitude levels for each stimulus were normalized so that differences in both the body and onset of each tone were less than 0.5dB, just beneath JND [just noticeable difference]” (p. 31). While a first experiment tested only brass instrument majors (n=31), using non-normalized stimuli, a second experiment using non-brass instrument majors (n= 32) produced the result mentioned above (p. 31). Trying to identify possible cues facilitating discrimination, Diaz reports that the process of amplitude normalization “also changed measures of spectral centroid between the tones” (p. 33); spectral centroid is a measure identifying the center of mass in a spectrum and has been linked to the percept of brightness (cf. Schubert, Wolfe, & Tarnopolsky, 2004). Specifically, the order of stimuli ‘intensity’ based on centroid values was ‘pah’, ‘tah’ and ‘dah’, with ‘pah’ being highest and ‘tah’ lowest. Very little research has been conducted on measures of spectral centroid[38], so it was not known whether these levels of difference would be noticeable by listeners” (p. 33). Spiegelberg (2002) also investigated the (production and) perception of articulation in music, with a focus on trumpet players. His pilot study showed that “trumpet attacks were distinguished by amplitude and length, whereas violin attacks were distinguished by spectral centroid, asynchrony, and maximum amplitude” (p. 131); while this finding might indicate that trumpet articulation has simpler acoustic parameters than violin articulation, rendering perception easier, his main study using only trumpet tones qualified this result. “This experiment revealed that segment length, centroid, and amplitude are key factors in listeners’ determinations of articulation similarity, though not the only factors” (p. 131). Interestingly, a perception experiment completed by 175 music and non-music majors showed that “[l]isteners can perceive four or five different types of articulation” (p. 131) even when the performer tried to distinguish a larger number of attack qualities. In line with results reviewed above,

[38] Kendall and Carterette (1996) report that their earlier research confirmed “the mapping of *nasality* … to spectral centroid” (p. 97; emphasis in the original; cf. Kendall & Carterette, 1989, 1993a & 1993b), and has been called “centroid *spectral brightness*” by some researchers (Kendall & Carterette 1996, p. 97; emphasis in the original).

> trumpeters were the best at this task [distinguishing trumpet articulations], followed by the rest of the brass players with the woodwinds and percussionists. As percussionists are used to listening to aspects of sound other than pitch, it is not surprising that they did well in this test (p. 202)[39].

Returning to the matter of the just noticeable difference in regard to spectral centroid, Carral (2011) published a paper in the same year as Diaz's article, which attempted to address precisely this question. She used spectral morphing (interpolations to gradually reduce the differences in spectral content) between two trombone sounds played with a modern and a historical French mouthpiece to determine the Just Noticeable Difference (JND, also called 'Difference limen' or differential threshold; Moore, 2003, p. 401); the two notes of the same pitch (Bb3) were produced by a semi-professional player in an anechoic chamber (p. 467). 26 students of Physics (p. 472) completed a two-alternative forced choice psycho-acoustical test using thirty intermediate steps and preserving the amplitude and frequency variations of the original sound. The results showed a JND of 0.86dB at 50 percent accuracy, and 1.28dB at 75 percent accuracy. "Likewise, the JND of the mean of the absolute difference in normalised spectral centroid" was determined at 0.08 and 0.11, respectively (p. 475). This finding agrees with similar results reported by Kendall and Carterette (1996; 0.11 for musicians and 0.15 for non-musicians), providing "evidence to support the assumption that the JND they reported are general and that they can be applied to other harmonic sounds" (Carral, 2011, p. 475)[40].

In summary, this section suggests that brass players have the ability to detect subtle differences in the sound produced by players from different language groups, possibly even more so in the case of the trombone than for other, higher-pitched brass instruments.

2.7 Previous attempts at investigating language influence on brass playing

Two Doctor of Musical Arts dissertations conducted at American universities have investigated the issue of language influence on brass playing. The more comprehensive of the two is Cox's D.M.A. dissertation (2014), in which the author

[39] Note, however, that the percussionist's perceptual abilities would be predicted by Schneider et al.'s (2005b) findings, albeit based on grounds that differ from Spiegelberg's reasoning.

[40] See Schubert, Wolfe & Tarnopolsky (2004) for an experiment on the perception of "[s]pectral centroid and timbre in complex, multiple instrumental textures" (title of the paper), and a discussion on how f0 affects the perception of 'brightness'.

calculated average formants of the vowels used in speech (F1 and F2) and the timbre produced during trombone playing (the second and third harmonics, corresponding to F1 and F2 in speech) for eighteen British (BE) versus twelve AE-speaking amateur trombonists. The author found differences at the large group (BE versus AE), dialect subgroup (different dialects of BE) and individual levels, although no statistical tests were performed to determine whether these differences were significant (visual inspection of the graphs included in the dissertation suggests they are). Interestingly, participants in this study were not able to correctly identify the nationality of professional players on different recordings of the same piece of music but some participants performed better than chance when asked which rendition they preferred, more often than not selecting recordings of players from their own language background.

A second study that used a combination of qualitative and quantitative data is Mounger's D.M.A. dissertation (2012), in which the author aimed to analyze "the orchestral trombone sound of France, Germany, and the United States … through the lens of language" (p. ii). The dissertation includes an analysis of "representative audio recordings of trombonists from France, Germany, and the United States" and tries to explain its impressionistic account of differences in sound by looking at the frequency of each vowel in the five hundred most common words in these languages. While the author tried to describe the trombone playing on selected recordings (including both solo and orchestra settings) "as succinctly and in the most objective terms possible" (p. 82), several problems exist with the qualitative part of her research, not least that a native speaker of Italian was included as an exponent of American playing based on the fact that he is the principal trombonist of a renowned American Orchestra.

Underlying the quantitative part of this and Cox's study seems to be the widely-held, but empirically questionable belief that brass players use the full range of vowel tongue positions during playing. Both authors certainly are not alone in making this assumption as shown by the survey of the brass pedagogical literature above; even if brass players used a wider range of tongue positions than empirical research attests, simply looking at the relative frequency of the tongue positions correlated with these sounds across different languages could provide only a weak chain of evidence. Thus, while Cox seems to have been able to uncover some evidence for the hypothesis of language influence (although lacking statistical validation), Mounger writes in her conclusion that "the quantitative data for this research does not fully support the concept that language can affect one's natural sound production on the instrument" (p. 83). I contend that both studies leave considerable room for improvement in terms of statistical validation, and participant classification and linguistic analysis, respectively.

Another study that explicitly mentions the possibility of L1 influence on brass playing is Budde's Ph.D. thesis (2011) on *Methods for Teaching Middle School Band Students to Articulate.* The author provides a comprehensive survey of the syllables that have traditionally been used to teach articulation on wind instruments and subsequently provides a careful account of the possible phonetic realizations of the consonants and vowels contained in these syllables within different languages. This leads him to state that "these differences manifest themselves in various ways as performers articulate on wind instruments" and to formulate the requirement that these "be taken into account when devising a method for teaching articulation to young musicians" (p. 5; see Torres, 2012 for a study on articulation on the flute). No evaluation of language background is included in the comparison of students assigned to certain treatment groups, however, other than creating an 'articulation guide group' which received and regularly reviewed an 'articulation guide sheet' developed by the author "based on the study of phonetics, native language, and music pedagogy" (p. 238). Students in this group scored significantly lower in the final evaluation of their ability to "articulate[d] clearly with accurate execution across various tempos" (p. 219) than students in the other test conditions, including a "practice group" and an "audio model group."

2.8 Chapter summary

This chapter has provided the reader with a basic understanding of brass instrument acoustics and more complex acoustical considerations underlying the influence of vocal tract resonances on brass instrument sound. While there is currently no conclusive evidence for such a phenomenon, the underlying physics suggest that vocal tract influence on brass instrument sound should be possible, which forms an important prerequisite for language influence on brass playing. After this, an overview of historical and contemporary method books was presented, which documented the long history of the use of speech syllables in brass pedagogy, outlining another important link between speech production and brass instrument performance. A comprehensive account of previous empirical research on brass playing was then presented, describing various aspects of vocal tract behavior during brass playing. The chapter concluded with sections on the perception of brass instrument sound by (expert) listeners, and a summary of two previous studies investigating language influence on brass playing.

3 Physiology and motor control of the upper vocal tract during speech production and brass playing

This chapter provides a brief overview of the physiology of two important articulators involved in speech production and brass playing, the jaw and the tongue. Specifically, it is shown that the jaw's function is mostly co-articulatory, whereas the tongue is a highly-flexible articulator that features independent control of at least two functional sections. Controlling such an articulator with an almost unlimited number of degrees-of-freedom is computationally expensive, calling for motor control models that can reduce such complexity. A number of different concepts in human motor control will be discussed, along with their application to speech motor control and observations concerning the influence of individual differences in vocal tract morphology and biomechanics on speech production. Subsequently, a discussion will be presented on whether motor control is shared across speech production and other vocal tract behaviors, followed by a comparison of experimental findings from studies reviewed in previous sections that document the role of certain articulators during speech production and brass playing. A focus will be placed on modular accounts of motor control that posit functional modules or 'muscle synergies' defined by patterns of muscle activations grouped according to their function (cf. Bernstein, 1967). Such accounts allow for cross-system interactions (cf. McFarland & Tremblay, 2006), and provide principled predictions for the role of language influence on brass playing.

3.1 Upper vocal tract physiology and articulator movement

Although fulfilling important functions during both speech production and brass playing, the lower parts of the facial-oral-laryngeal-respiratory musculature (see Gick, Wilson, & Derrick, 2013) will be disregarded for the purposes of this overview, and I will focus instead on those areas that can be observed during (or influence) the acquisition of ultrasound data of the tongue. More specifically, the focus will be on the tongue and movements of the jaw. Although these cannot be measured using the experimental setup employed in this thesis, jaw movements are a pivotal factor to account for in regard to the application of a jaw brace that ties tongue motion to jaw motion, and the underlying dynamics of tongue-jaw coarticulation. Other structures not discussed in this section include the role of the glottis and pharyngeal constrictions during brass playing; these have been described in the review of empirical brass research above, and will be revisited in light of my findings in the discussion section of this thesis.

3.1.1 The jaw

The main function of the human jaw (also jawbone or mandible) is mastication (or chewing), for which it is outfitted with two sets of powerful shallow muscles: the masseter ("chewer") and temporalis muscles (Gick, Wilson, & Derrick 2013, pp. 148-149). The subtle movements essential for speech production, however, are realized by a number of deep muscles attached to the inside of the mandible. While superficially the jaw might seem to operate much like a hinge, the temporomandibular joint (which connects the mandible to the temporal bone of the skull) is not fixed in place, allowing the mandible to slip out of this joint and move in any direction. It thus has the ability to use the maximum possible number of degrees-of-freedom (Gick, Wilson, & Derrick, 2013, pp. 147-148). Movements off the midsagittal plane, however, are not thought to be of "primary importance for speech," though they are very useful for grinding food (Gick, Wilson, & Derrick, 2013, pp. 151, 148). Lowering the mandible with precision is accomplished by a number of muscles connecting the underside of the jaw to the hyoid bone, which in turn is connected to the larynx via muscle bands figuring prominently into determining its vertical position. These muscles referred to as jaw depressors also contribute directly to motions of the tongue body due to their ability to tense "the floor of the mouth on which the tongue rests" (Gick, Wilson, & Derrick, 2013, p. 152). Some of the muscles described above are illustrated in figure 3.1 on the following page.

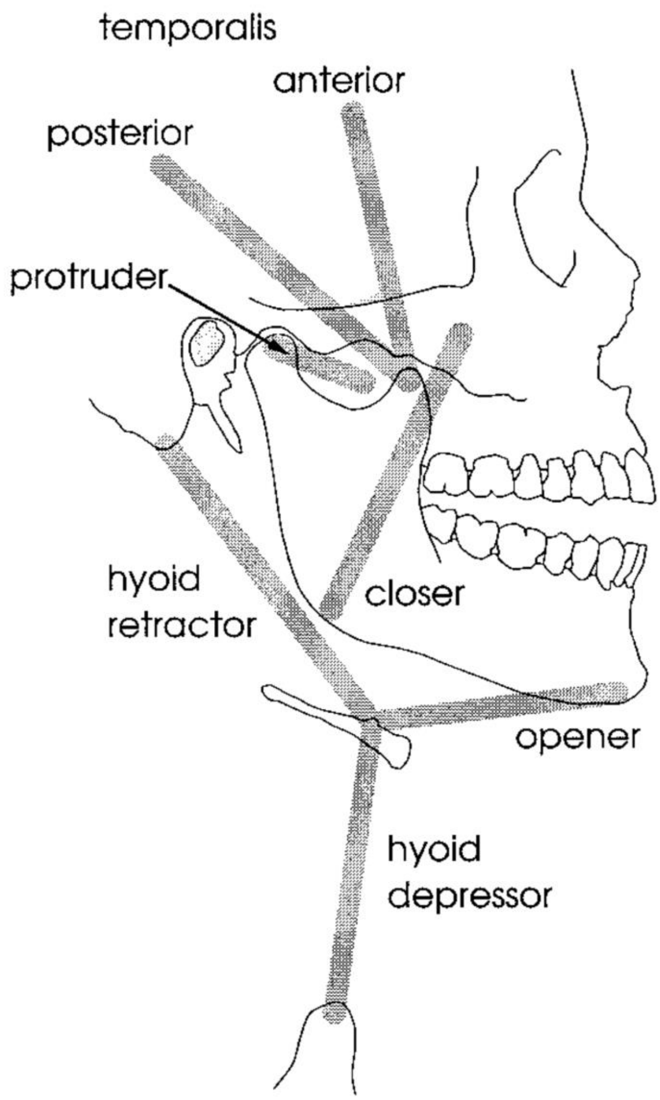


Figure 3.1: Schematic of the modeled muscle groups and their attachments to the jaw and hyoid bone. Reproduced with permission from Shiller, Ostry, & Gribble, 1999, p. 9074.

3.1.2 The role of the jaw during speech production

The jaw has been the subject of a large number of studies on speech production and motor control, even though it is thought to be solely responsible for phonemic contrast in only one language; this is the /s/ - /θ̠/ contrast in Icelandic (Pétursson, 1971; note that this finding, however, is based on a single speaker only).

Amerman, Daniloff and Moll (1970) were among an early group of researchers who used x-ray imaging to investigate the movements of the jaw during speech production. Observing the effect of the vowel /æ/ on one or more preceding consonants produced by four speakers of American English, their results showed that jaw displacement could affect up to three consonants preceding the vowel target, and that measurements of the amount of “jaw opening achieved during running speech were uniformly smaller” than during sustained production of /æ/ (p. 147). A later study on English and Swedish by Keating et al. (1994) investigated jaw coarticulation across various vowel and consonant contexts, demonstrating that vowels are more likely to influence jaw opening during consonant production than vice versa, and that average jaw opening patterns with vowel height; note, however, that this relationship is not linear, with /e/ being disproportionally open (p. 413). Regarding variability, these authors’ results seem to support a finding by Crystal and House (1988) that longer segment duration leads to increased jaw opening and variability (vowels were more open than consonants “by about 3 mm [on] average”; Keating et al., 1994, p. 421), and the findings of other previous research whereby “lower vowels vary more than higher ones” (Keating et al., 1994, p. 420; cf. Tuller, Harris & Gross, 1981; Imagawa et al., 1985, Farnetani & Faber, 1992). Regarding consonants, the results of this and other studies suggest that alveolars are most resistant to coarticulation, varying least across different vowel contexts, with velars being much more affected (Keating et al., 1994, p. 420; cf. Tuller, Harris, & Gross, 1981; Kiritani et al., 1983; Jun et al., 1991). In all of these cases, jaw movement seems to interact with tongue height to produce various constrictions in the oral cavity; it can, however, also interact with tongue position in the horizontal dimension (along the occlusal plane), and differences in the amount of vertical and horizontal tongue/jaw coordination are thought to be influenced by “individual anatomical structures as the palate of every speaker differs in length and steepness of the area between alveolar ridge and palatal vault” (Kühnert et al., 1991, p. 24). Individual differences affecting speech production will be discussed in more detail in section 3.3.4 below.

3.1.3 The tongue

The human tongue has often been compared to an elephant’s trunk or the tentacles of an octopus, due to its seemingly unrestricted movement potential. While we have since learned that octopuses create ‘pseudo’ joints (proximal, medial and distal) in each tentacle to simplify the degrees-of-freedom problem (Sumbre et al., 2006; cf. Gick, Wilson, & Derrick, 2013, p. 144), the same strategy does not seem to be applicable to the human tongue, possibly because it is operating in a highly constrained space (Gérard et al., 2003/2006, p. 3). Recent research, however, has made great strides towards a more accurate understanding of the human tongue by leaving behind earlier two-dimensional models (cf. Gick, Wilson, & Derrick, 2013, p. 144) and instead modelling tongue behavior in three-dimensional space (Dang & Honda, 2002; Napadow, Kamm, & Gilbert, 2002; Gérard et al., 2003/2006; Vogt et al., 2006; Buchaillard, Perrier & Payan, 2009; Fang et al., 2009; Stavness, Lloyd, & Fels, 2012). A central assumption underpinning this research is the concept of the tongue as a muscular hydrostat, meaning it preserves its volume whichever shape it may take; furthermore, its biomechanics bear more resemblance to hydraulic systems than "the mechanical lever arms used by most skeletal muscles" (Sanders & Mu, 2013, p. 1103). Based on empirical accounts of observed tongue motion and/or anatomy, all of these models divide the tongue into at least two functionally independent sections (MacNeilage & Sholes, 1964; Öhman, 1967; Smith, 1971; Miyawaki et al., 1975; Stone, Epstein, & Iskarous, 2004). Before discussing the evidence supporting this assumption in more detail, it is important to first recognize the tongue as a collection of individual muscles.

There are four intrinsic muscles within the tongue "that can squeeze in different directions" (Gick, Wilson, & Derrick, 2013, p. 168), illustrated in figure 3.2 on the next page. They can be thought of as pairs that allow humans, (1) to narrow/widen the tongue (transversus and verticalis muscles; these muscles can also effect elongation and thickening/flattening of the tongue; Gick, Wilson, & Derrick, 2013, p. 169), and (2) to curl the tip and sides of the anterior tongue upward/downward, giving the tongue an overall concave/convex shape (superior and inferior longitudinal muscles; Gick, Wilson, & Derrick, 2013, p. 170). These muscles can be identified as "discrete entities in some region of the tongue" but mostly separate into "smaller muscle fascicles that interweave with … other tongue muscles," rendering them morphologically indistinguishable (Sanders & Mu, 2013, pp. 1103-1104). Such muscular configuration is one reason why it is difficult to determine the exact workings of the tongue. In addition to these intrinsic muscles that can alter the shape of the tongue, several extrinsic muscles attach the tongue to the surrounding anatomical structures, causing it to move as a whole (cf. Sanders & Mu, 2013, pp. 1103) – these are shown in figure 3.3 below. The largest of these, the genioglossus muscle, "originates at the mental spine of the mandible" and "often acts like an intrinsic muscle, as it makes up more of the bulk of the tongue than any other single muscle" (Gick, Wilson, & Derrick, 2013, pp. 152-153). It has "several distinct regions that can be independently controlled" to protract and depress the tongue; it can also “help to create a midline groove in the tongue” (Gick, Wilson, & Derrick, 2013, p. 153). Other extrinsic muscles connect the tongue to the palate (palatoglossus), hyoid bone (hyoglossus), and the styloid processes (styloglossus), helping to elevate and retract, pull down- and backwards on (especially the dorsum), and stabilize the tongue, respectively (Gick, Wilson, & Derrick, 2013, p. 154).[41]

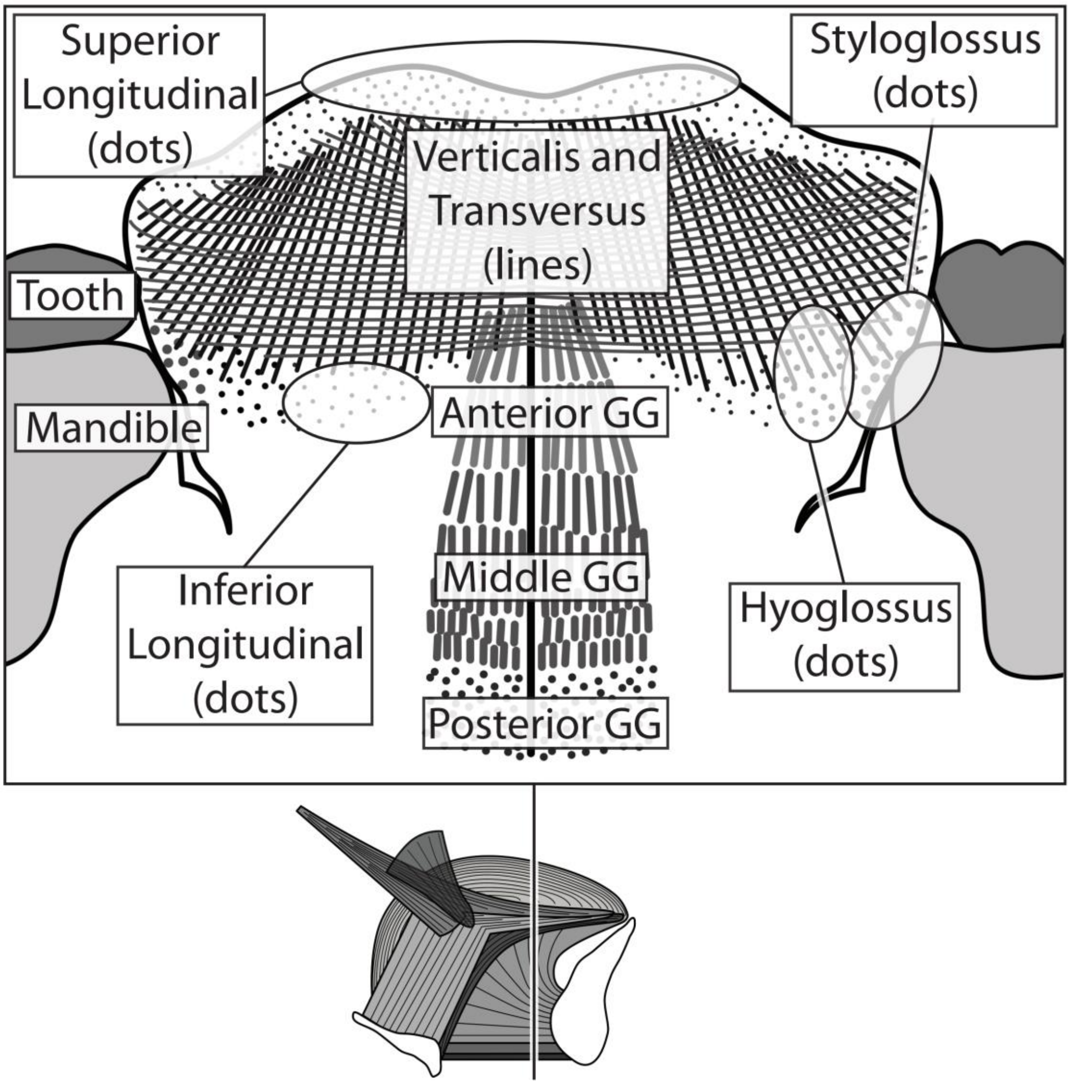


Figure 3.2: “Tongue muscles: coronal cross-section through the mid-tongue; the location of the cross-section is indicated by the vertical line through the sagittal tongue image at the bottom (image by D. Derrick, inspired by Strong, 1956).” Reproduced with permission from Gick, Wilson, & Derrick, 2013, p. 169. © 2013 Bryan Gick, Ian Wilson, and Donald Derrick.

[41] Older accounts thought the styloglossus helped raise and pull back the tongue; this notion was rectified by Takano and Honda (2007).

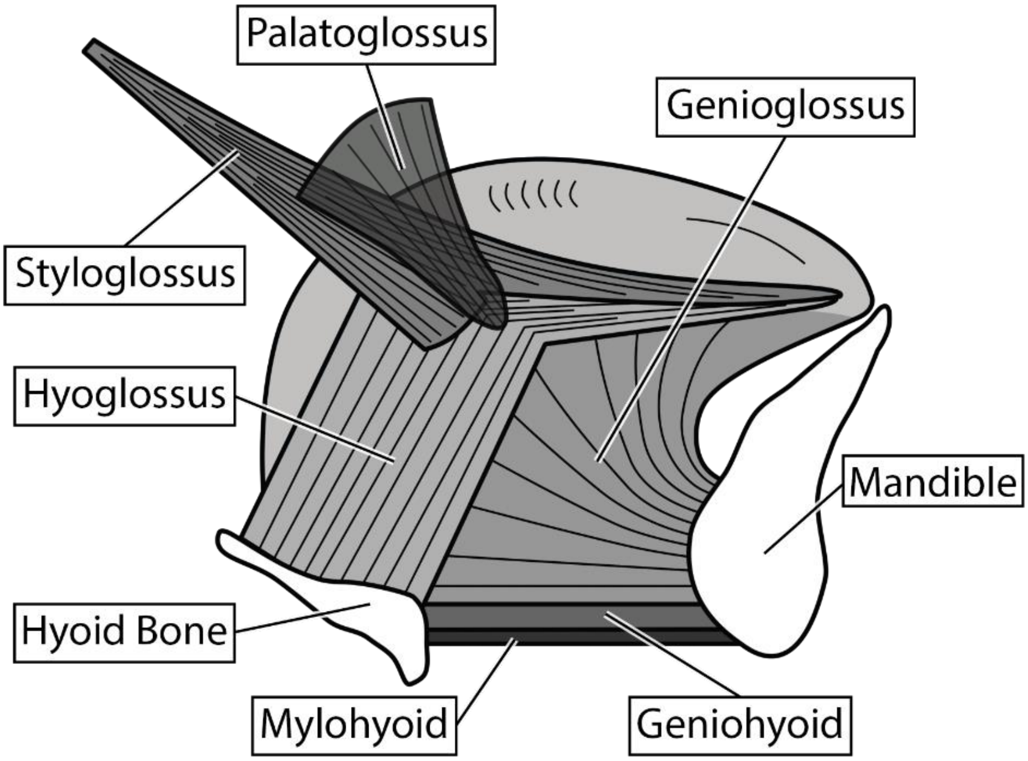


Figure 3.3: “Extrinsic tongue muscles: right side view. Geniohyoid and mylohyoid are included for the context only (image by D. Derrick).” Reproduced with permission from Gick, Wilson, & Derrick, 2013, p. 152. © 2013 Bryan Gick, Ian Wilson, and Donald Derrick.

While functional independence has long been assumed for various sections of the genioglossus muscle (cf. MacNeilage & Sholes, 1964; Harshman, Ladefoged, & Goldstein, 1977; Kakita, Fujimura, & Honda, 1985; among others), independent control of subsections of other intrinsic tongue muscles was only recently proposed and verified using ultrasound and tagged-MRI data in a paper by Stone, Epstein and Iskarous (2004). Specifically, this paper asked whether the tongue is composed of two parts (i.e., a body and tip), or multiple functional segments, and then used correlations between measurements of tongue movement at different sections of the vocal tract to demonstrate functional independence (p. 520). The argumentation presented in the paper is the following: “[h]igh correlations between tongue segments would suggest passive biomechanical constraints [while] low correlations would suggest active independent control” (p. 507). Interestingly, analysis of their data provides evidence for both “physiological *and* higher level linguistic constraints” (p. 507; italics added). Ultrasound data collected midsagittally and at five different coronal slices of the tongue supports “an overall division of the tongue into two large regions in which adjacent segments correlate positively, and distant ones negatively,” and also that “segments can be coupled for certain motions and that the correlations and couplings are phoneme dependent” (p. 515; cf. findings by Miyawaki et al., 1975 regarding the genioglossus muscle). These findings are further supported by measurements of tongue compression patterns using tagged-MRI; “[a]lthough tMRI cannot definitively distinguish active from passive [muscle] compression” (p. 517), the observed patterns suggest that, similar to the genioglossus, the transversus and verticalis muscles can be controlled in segments (anterior, posterior; pp. 518-519). Furthermore, the measurements indicate that the genioglossus and transversus muscles are often coupled, while the verticalis muscle can be controlled independently; “one-sided or weighted activation patterns” of this muscle show how it can be used “to create left-to-right asymmetries” (p. 519).

Being a complex muscular structure, the tongue requires a lot of dedicated brainpower to operate, and is disproportionally represented on both the sensory and motor homunculi indicating the mapping of body parts onto the sensory and motor cortices in the brain (Gick, Wilson, & Derrick, 2013, p. 25, cf. Penfield and Roberts, 1959, p. 27). Note that the substantial representation of tongue on the sensory cortex does not automatically imply conscious awareness of the tongue’s state; in fact, proprioception may be relatively limited in the tongue (Gick, Wilson, & Derrick, 2013, p. 173). Although “proprioceptive afferents” have been shown to exist in the human tongue (Adatia & Gehring, 1971), contrary to an early account questioning their existence altogether (Carleton, 1938), the notion of them carrying primary importance for conscious proprioception or the ‘position sense’ was regarded to be widely dispelled by Matthews (1964). Rather, Matthews stated, the role of such muscle spindles “is now usually believed to be to play a part in the subconscious nervous control of muscular contraction, both during movement and during steady contraction” (p. 273). Such a property still “provide[s] the nervous system with information about the state of the muscles,” though not at a conscious level (p. 273). The most recent findings on the topic, however, seem to show that the true extent of proprioceptive ability in the tongue is yet to be determined. Trulsson and Essick (1997) report that “within the lingual nerve of humans are found deep mechanoreceptors that possess the capacity to signal tongue movements in the absence of direct contact with the receptive field” (p. 746). These deep receptors would provide proprioceptive feedback in the classical sense, while superficial mechanoreceptors lacked this capacity (p. 746). The proximity of the tongue to several rigid structures, however, might provide additional information via the spatial and temporal patterns of contact-evoked discharges in the superficial mechanoreceptors (p. 746). This account is consistent with Carleton's (1938) interpretation of her early findings (see above) "that conscious tongue position sense (kinesthesis) [is] normally mediated by spatial cues provided by contact of the lingual mucosa with stable intraoral landmarks" (Trulsson & Essick, 1997, p. 746). Some indication of why previous studies may have produced seemingly contradictory results comes from further research conducted in the meantime (Siegel & Hanlon, 1983; Grover & Craske, 1991). Trulsson and Essick interpret these findings to indicate "that central neural mechanisms underlying perception process information from the superficial mechanoreceptors and from the deep mechanoreceptors independently" (p. 746). Benedetti (1988) showed "that proprioceptive signals from the tongue muscles do not serve to reorient the mucosa's spatial frame of reference in response to physiological perturbations" (Trulsson & Essick, 1997, p. 746), which can provide an explanation for different effects of deafferentation, depending on the tongue's position when subjected to sensory stimuli.

The extensive representation of the tongue on the motor cortex is also indicative of its high number of degrees-of-freedom; a possible explanation for how the human brain manages to reduce these degrees-of-freedom and cognitive load (cf. Bernstein, 1967; cf. Gick & Stavness, 2013; Gick, 2016) may be offered by accounts of tongue 'bracing' against a "rigid vocal tract surface, such as the teeth or palate" (Gick et al., in press, pp. 3, 24-25; cf. Stone, 1990). In a forthcoming paper, Gick et al. (in press) used electropalatographic (EPG) data and biomechanical simulations to show that the tongue remains in "almost constant contact with fixed vocal tract surfaces, either laterally or coronally, with the notable exception of a subset of instances of low vowels and postvocalic /l/" (p. 14). Most commonly, bracing happens laterally (this was observed for "97.5% of the total duration of the recordings" for two speakers contained in the EPG database used by Gick et al., p. 9), although it often occurs coronally during laterals, and "the dorsal tongue increases contact with lateral pharyngeal structures … during retracted postures for low vowels" (p. 16). The biomechanical simulations carried out for the paper additionally show that "bracing requires active muscular control" (p. 21), a conclusion which is supported by the EPG finding "that coronal bracing is always in place before lateral bracing is lost" (p. 22). Bracing does not need to be bilateral: "both speakers [in the EPG experiment] showed an asymmetrical tendency to release contact on one side of the mouth versus the other side" (p. 10), an asymmetry which had previously been documented for the phonemes [s], [t], and [k] in English by McAuliffe, Ward, & Murdoch (2001), and which might interact with handedness (Gick et al., in press, p. 23; cf. also the findings by Stone, Epstein, & Iskarous, 2004, discussed above). In addition to reducing the possible degrees-of-freedom by stabilizing the tongue (for which unilateral bracing might suffice; p. 9), Gick et al. also suggest that lateral bracing forms a requirement for many speech sounds by creating a "closed aeroacoustic tube that directs the airstream through any medial speech constriction" (p. 4; cf. Perkell, 1979; Honda, Takano, & Takemoto (2010).

3.2 Human motor control

3.2.1 Models of human motor control

While the above overview of two important articulators involved in both speech production and brass playing has provided information on some commonly observed and potential movement patterns of these articulators, no information was provided on how these processes are controlled. A major issue in human motor control concerns the question of how the central nervous system (CNS), comprised of brain and spinal cord, manages to set complex combinations of muscles in motion to achieve various motor goals. Prevalent approaches postulate 'open-loop'[42] and 'closed-loop' models of motor control, which differ regarding the role of feedback (sensory or 'afferent' information; cf. Schmidt & Lee, 2011, p. 135). "A closed-loop system depends heavily on the involvement of particular types of sensory information as it executes its function" (Schmidt & Lee, 2011, p. 135), while in the case of an open-loop system, certain instructions are sent to an effector that "carries them out without the possibility of modification if something goes wrong" (Schmidt & Lee, 2011, p. 178). These concepts are not necessarily mutually exclusive, as they could apply to "fundamentally different" kinds of both "fast and slow" motor movements (Schmidt & Lee, 2011, p. 161). Closed-loop systems carry considerable explanatory power for many motor actions, "especially those that require a system to 'control itself' for long periods of time" (Schmidt & Lee, 2011, p. 135). They are, however, incompatible with rapid

[42] The term 'open-loop' system is used in reference to a "feedback pathway [that] is 'cut' or 'open,' as when a switch is open" (Schmidt & Lee 2011, pp. 177-178).

movements, as there would not be “sufficient time for the system to (a) generate an error, (b) detect the error, (c) determine the correction, (d) initiate the correction, and (e) correct the movement before a rapid movement is completed” (Schmidt & Lee, 2011, p. 161).

Studies trying to determine whether open- or closed-loop systems more appropriately describe certain movements, have proposed at least two levels of control within the motor system,

> (a) an executive level (including information-processing stages) for selecting, organizing, and preparing and initiating a complex pattern of muscular activities, and (b) an effector level (motor programs and the muscular system) for actually controlling or producing the patterns as they unfold (Schmidt & Lee, 2011, pp. 196, 198).

This division seems to be supported by studies on ‘startled actions,’ during which the presentation of a startle stimulus (“a very loud (130 dB) auditory signal”) results in an “involuntary initiation” that otherwise would have occurred in response to the regular (visual) stimulus (Schmidt & Lee, 2011, p. 192). Open-loop models of motor control differ from closed-loop models at the effector level (b), so it is therefore possible that the execution of a preplanned movement in response to a startle stimulus could bypass the executive level[43]. Reaction times in a study by Wadman et al. (1979) on the control of “fast goal-directed arm movements” (p. 3) were different for normal trials and those initiated by startle stimuli of various intensities (Schmidt & Lee, 2011, p. 192). Otherwise, “all of the movement events, including magnitude and timing of the EMG patterns, were identical to those in the normal trials,” while accompanied by “the typical startle response (eye blink and neck reflexes)” (Schmidt & Lee, 2011, p. 192; cf. Carlsen, Maslovat, & Franks, 2012). The difference between averaged premotor reaction times was nearly 100 ms when comparing normal trials to startle trials (171 ms versus 77 ms, respectively). Reaction time (RT) is used frequently as a measure of complexity for motor tasks, with “longer-duration (but still rapid) movements produc[ing] longer RTs” (Schmidt & Lee, 2011, p. 196; cf. Klapp, 1996; for an application of this measure to speech production see Klapp, 2003; Krivokapic, 2012).

[43] Stevenson et al. (2014) report disagreement among researchers regarding “the neural structures involved in this initiation process” (p. 30); while some researchers believe that a startle “causes release of the stored response through subcortical structures” (e.g., Valls-Solé et al., 1999), others have postulated the existence of a subcortical pathway triggering movement (e.g., Alibiglou & MacKinnon, 2012).

However, the above experiment reports no clear differences in the way of complexity and duration between normal- and startle-trials (Schmidt & Lee, 2011, p. 192).

3.2.2 Generalized motor programs

Apart from closed-loop and open-loop models of motor control, some researchers have assumed the existence of generalized motor programs, referring to “a particular *class* of actions … stored in memory” which would result in “a unique pattern of activity” when executed (Schmidt & Lee, 2011, p. 208; emphasis in the original). For example, the underlying program is understood to be generalized in the sense that choices when setting the values of various parameters would “alter the output” so that the same motor program could be used to perform similar movements with different limbs (Schmidt & Lee, 2011, p. 208). Evaluating the accuracy of these predictions has been compromised, however, by invariance observed in different motor actions, especially in relation to their relative timing; it is not clear whether such invariance has its origins in the central mechanism that “triggers an effector into action” or “the (somewhat variable) motor delays (such as neural delays and muscle recruitment time) that occur following a central trigger” (Schmidt & Lee, 2011, p. 220; cf. Heuer, 1991). Furthermore, studies investigating motor transfer from one task to another have found it to be “typically small,” even for similar tasks, and these results agree with findings “that motor abilities are both numerous and specific, and that even similar tasks appear to correlate very weakly with each other (with the possible exception of timing skills)” (Schmidt & Lee, 2011, p. 482). Nonetheless, there could be both positive and negative transfer across two tasks containing similar elements (Schmidt & Lee, 2011, p. 485), which has been frequently claimed to hinder Second Language acquisition.

In opposition to generalized motor control, Henry (1958/1968) argued that motor abilities are “*specific* to a particular task” – unsurprisingly, this was a controversial idea at the time (Schmidt & Lee, 2011, p. 304); this ‘specificity hypothesis’ has now been “widely accepted for many years” (Schmidt & Lee, 2011, p. 305). “[A] general ‘timekeeping’ ability” may, however, underlie performance of some tasks (Schmidt & Lee, 2011, p. 305), as suggested by research focusing on “how the temporal aspects of movements are organized in the central nervous system” (Schmidt & Lee, 2011, p. 305; cf. Keele & Hawkins, 1982; Keele, Ivry, & Pokorny, 1987). This research conducted by Keele, Ivry and colleagues includes a study by Keele et al. (1985) that found relatively high correlations between perception and production of timing, suggesting a link between the two; a speech perception and production task formed a part of that study, and several studies have since shown a close link between speech perception and production (Perkell et al., 2004 & 2006; Beddor, 2015; among others). These findings suggest that there should be very limited transfer of motor control from speech production to vocal tract control during brass playing, if at all. More recent research reviewed in the following two sections, however, suggests otherwise, with different conclusions potentially explained by the higher amount of movement details observed in recent studies.

3.2.3 Modular accounts of motor control

An explanation for why motor control should be understood as neither completely generalized nor completely specific to any particular task arises from recent experimental findings supporting an account of motor control as modular. 'Modular' as used in this research refers to functional modules as "discrete entit[ies] whose function is separable from those of other modules (Hartwell et al., 1999, p. C48)[44]. In a recent review paper, d'Avella et al. (2015) highlight modularity as "the key organizational principle that the central nervous system employs for achieving versatility and adaptability in motor control" (p. 1). A related, but more closely defined concept is the notion of 'muscle synergies,' which refers to the modular organization of multi-muscle activity across different motor tasks, and which is thought to be encoded in the CNS (d'Avella et al., 2015, p. 2; cf. Bizzi & Cheung, 2013). Increasing evidence points to the existence of such muscle synergies not only in "mature, fully developed systems" but also to their role in evolutionary and developmental processes (Lacquaniti et al., 2013).

Dominici et al. (2011) show that primitive, inborn patterns of locomotion are retained and tuned throughout life, while new patterns are added during development (p. 997). Their findings are derived from motor neuron activity of multi-muscle recordings in stepping newborns, toddlers, preschoolers, and adults; "markedly similar patterns" were also found in the rat, cat, macaque, and guineafowl, suggesting that locomotion in several animal species is built upon common primitives (p. 997; cf. Lacquaniti et al.,

[44] Note that a similar concept was referred to as 'coordinative structures' by Turvey and colleagues (Turvey, 1977; Turvey, Shaw & Mace, 1978), based on Easton (1972). Later studies, however, moved away from neurophysiology, thus modifying the original meaning of the term (e.g., Kelso et al., 1986; cf. Gick & Stavness, 2013).

2013, p. 5). In terms of life-span development, these findings “strongly suggest a continuous development of the corresponding motor modules” (Dominici et al., 2011, p. 998), which corresponds well with individual differences in motor control that could be “due to the influence of training and experience” (Torres-Oviedo & Ting, 2010, p. 3093). These differences, in turn, could have their origin in the “range of behaviors that are used to tailor the specific structure of muscle synergies in an individual” (Torres-Oviedo & Ting, 2010, p. 3093). Bizzi and Cheung (2013) also mention the possibility of new muscle synergies resulting from highly skilled movements such as those required for piano playing or professional sports (p. 2). A related question concerns the question of whether muscle synergies (or modules) are shared across tasks or are rather specific to a certain task. A paper by Chvatal and Ting (2013) found both shared and task-specific muscle synergies for participants’ responses to perturbations during standing, walking, and for unperturbed walking. Likewise, Frère and Hug (2012) were able to demonstrate that muscle synergies are shared by experienced gymnasts for a skilled motor task that requires learning.

3.2.4 Is motor control optimal?

The fact that similar muscle synergies were found across participants in the study by Frère and Hug (2012) above seems to suggest that an optimal solution exists for any given motor task, and that any individual should arrive at this solution with sufficient training (and in absence of other limiting factors). Indeed, for many years, research on motor control has assumed that the optimization of motor behavior represents the central goal of human motor planning. A recent paper by Ganesh et al. (2010), however, suggests that “the search for optimality could be secondary in motor planning” (Perrier, 2012, p. 230; cf. Friston, 2011). Using a complex motor task (“maintaining the wrist within a certain angle amplitude under the influence of vibratory disturbances”), EMG measurements of the involved muscles showed that subjects “(1) did not all use the same motor control strategies to compensate for the perturbation, and (2) did not in general adopt any optimal strategy even if they had previously realized the optimal movement” (Perrier, 2012, p. 230). The authors of the paper interpret these findings to signal that the observed motor behavior “is largely influenced by motor memory,” suggesting that “the CNS does not optimize co-activation tasks globally but determines the motor behavior in a tradeoff of motor memory, error, and effort minimization” (Ganesh et al., 2010, p. 382). Note that ‘motor memory' (alternatively 'muscle memory') is generally defined as "the persistence of the acquired capability for performance" (Schmidt & Lee, 2011, p. 461), while the exact nature of the concept could refer to a "motor program, a reference of correctness, a schema, or an intrinsic coordination pattern (Schmidt & Lee, 2011, p. 462; cf. Amazeen, 2002).

An argumentation similar to that put forward by Ganesh et al. (2010) is presented in Loeb (2012), who poses the question of whether it is appropriate to apply the theory of optimal control to biological organisms, and answering in the negative. Rather, the motivation for novel forms of behavior would usually be a 'good-enough' solution that does not entail the optimization of "a single, known cost function to be optimized" (p. 757). Trial-and-error learning employs such a good-enough strategy to find useful but not necessarily optimal patterns of behavior, leading to a diversity of solutions that offers robustness for the individual organism and its evolution (p. 757). Any potential solution or stable state depends on the "starting state and nature of the system," where the cost function would have many peaks and valleys, depending on the system's complexity (p. 758-759); likewise, the most effective exploration strategy would depend both on the system *and* the goals, strengths and weaknesses of its controller (p. 759).

3.3 Motor control during speech production

A central issue for researchers working on speech motor control concerns its frame of reference or the target domain of its movements. Much evidence points to the fact that speech motor goals are "defined in an abstract domain," given that "there is no unique physical correlate for a given elementary speech sound" (Perrier, 2006, p. 14); they can alternatively be regarded as situated either in articulatory, acoustic or perceptual space, and "a large variability of patterns" has been observed in the "neurophysiological, articulatory, and acoustic domains" (cf. Perkell & Klatt, 1986). It is also possible that speech production targets operate in "a multimodal space associating orosensory, auditory, and even visual characterizations" (Perrier, 2006, p. 14; cf. McGurk & MacDonald, 1976; Rosenblum, 2008; Gick & Derrick, 2009). This assumption is directly represented in the "many-to-one characteristic [of] the relationships between motor commands, articulatory positions, and acoustic or auditory properties" (Perrier, 2006, p. 14).

While similar challenges arise with the degrees-of-freedom problem in the context of, for example, limb movements, the various frames of reference further complicate this issue for the domain of speech production. Speech movements can be "very short, since vowels have a mean duration of approximately 80 ms and consonants have mean durations around 40 ms" (Perrier, 2006, pp. 14-15; cf. O'Shaughnessy, 1981). All these factors constrain the applicability of the various motor control models (described in section 3.2.1) to speech movements, especially when considering the role of feedback. Long-latency feedback requiring cortical processing does not seem to be usable by the speech production apparatus, given the brevity of speech movements. Similarly, auditory feedback would only be useful at a "suprasegmental level" and as part of an "*a posteriori* monitoring … to correct segmental aspects of speech after it was produced" (Perrier, 2006, p. 15; cf. Perkell et al., 1997). That intelligible speech is nonetheless possible without "online use of auditory feedback to control speech at a segmental level" has been shown through experimental work involving speakers suffering from hearing loss (Perrier, 2006, p. 15; cf. Lane & Wozniak, 1991; Nasir & Ostry, 2008).

Evidence indicating a reduction in the complexity of speech motor commands arises from 'quantal' relations that have been argued to underlie various acoustic-articulatory or auditory-acoustic relationships in speech production (Stevens, 1989; Stevens & Keyser, 2010). The use of the adjective 'quantal' in this context refers to a mapping that varies in steps rather than continuously, or as Stevens (1989) puts it: "there appear to be ranges of the articulatory parameter and other changes where the acoustic parameter is more sensitive to changes in articulation. The articulatory-acoustic relations are quantal in the sense that the acoustic patterns shows [sic] a change from one state to another as the articulatory parameter is varied through a range of values" (Stevens, 1989, p. 3). Recently, such quantal features have also been shown to exist in the biomechanics of the lips and larynx (Moisik & Gick, 2013; Gick et al., 2011; cf. also Gick & Stavness, 2013; Gick et al., 2014; Gick, 2016), meaning that "different initial muscle settings produce regions in which large variations in input activation yield stably different" states of the relevant articulators (Gick et al., 2011; 179). Such stable states seem to exist in terms of constriction degree at the lips (Gick et al., 2011, p. 179), and regarding the narrowing of the epilarynx at two different locations within the laryngeal region (Moisik & Gick, 2013). Regardless of at which level such a quantal relationship holds (i.e., biomechanics: muscle activation affecting articulator location, versus acoustics: articulator location resulting in a certain acoustic signal), it would seem to provide a greater margin for variability or error in terms of movement planning, while simultaneously reducing the need for continuous sensory and/or auditory feedback.

3.3.1 Models of speech motor control

Faced with various frames of reference that could serve as the target domain for speech production, a “large majority of speech motor control models published in the literature assume the existence of *internal representations* of the speech apparatus” (Perrier, 2006, p. 15; emphasis in the original; cf. Laboissière, Schwartz, & Bailly, 1991; Hirayama et al., 1992; Guenther, 1995; Perkell et al., 1997). One such concept is referred to as the ‘desired trajectory hypothesis’; in such a model, the CNS would come up internally with a number of “possible motor command sequences that would both generate the required trajectory and minimize a motor criterion,” and then work back in ‘inverse’ fashion to determine the relevant motor commands for the optimal motor sequence (Perrier, 2006, p. 16). A variation of this approach, called a ‘direct’ internal model, would make similar use of internal modeling to find an optimal motor solution; however, “the trajectory of the final effector toward the target would be a consequence of the planning strategy rather than the specification of the task itself” (Perrier, 2006, p. 16). More recently, the ‘optimal control feedback model’ proposed by Todorov and Jordan (2002) has provided a way of including online feedback in an internal control model, which would utilize feedback to “selectively modify motor commands in an optimal way when deviations in the task space occur that would endanger the achievement of the task goal” (Perrier, 2006, p. 16). To avoid the problems associated with delay of the traditional feedback channel, this model posits to treat the “outputs of the internal models as afferent [sensory] signals” (Perrier, 2006, p. 16).

While they were not developed specifically within the framework of speech motor control, Perrier (2006) attests that “internal models are very powerful tools for dealing with the many-to-one nature of the relations between motor commands and vocal tract configurations or spectral characteristics of the acoustic signal” (p. 16). They also offer a way of dealing with the long latency of feedback processed by the cortex, and the possible “multimodality of speech task representations” (p. 16). Experiments using perturbations “artificially introduced either in the speech production apparatus or in speech perception feedback” have resulted in compensation strategies that match up well with internal models of speech motor control (p. 17). Unfortunately, the predictive power of these internal models of speech production is diminished when applied to dynamic processes (p. 18; cf. Tremblay, Houle, & Ostry, 2008). This effect may be a result of “mechanical and dynamical properties of the peripheral speech apparatus” that vary during the movements, “since they depend on the positions of the articulators in relation to each other and on the strain of the soft tissues” (p. 19). In other cases, however, researchers may have tried too hard to model seemingly complex speech articulator trajectories that “could be simply due to biomechanical properties” of the articulators, as in the naturally occurring curved movement paths during VCV sequences involving velar stop consonants (Perrier et al., 2003; cf. Houde, 1968; Mooshammer et al., 1995).

Generalized models of motor control have also sometimes been applied to speech motor control, with results often failing to provide evidence for the existence of such models (Perrier, 2012, p. 220). Instead, “the focus of studies investigating the nature of dynamical internal models [has now] shifted from the investigation of their existence in the form of generalized models to the investigation of how multiple localized or specialized models could be learned and used to control different tasks in various contexts” (Perrier, 2012, p. 225).

It seems, then, that an accurate model of speech motor control needs to account for the different plausible frames of reference and the complex mappings arising from this situation, as well as allowing for a way to incorporate (limited) feedback resulting from a rapidly varying time domain. Perrier (2006) suggests that calculations of an ‘intermediate sliding variable’ (Slotine & Li, 1991) might provide a solution to this problem (p. 21). A sliding variable is defined by Hanneton et al. (1997) as “a specific combination of the instantaneous error and its successive time derivatives,” that allows the reduction of high-order control problems to a simpler calculation; this facilitates the use “both feedback and feedforward mechanisms to approximate the inverse transfer function of the controlled system” in real-time (p. 382; cf. d’Avella et al., 2015 for a related notion termed ‘intermittent control’). Speech movements controlled by such a mechanism would consist of an underlying low velocity gesture (under feedback control) and short superimposed ballistic submovements (under feedforward control) to correct for deviations from the estimations of the sliding variable (Perrier, 2006, p. 22).

Strong evidence for the reliance of speech production on feedforward control has recently been put forward through studies using an auditory startle paradigm to trigger speech movements. It has been hypothesized that the involuntary initiation of startled movements arises through activation at the effector level, ruling out the possibility of any sensory or auditory feedback (Schmidt & Lee, 2011, p. 192). In an initial investigation on the articulation of the syllable [ba], Stevenson et al. (2014) showed that "the kinematics and prepared syllable were initiated significantly earlier in startled trials," but the timing of "kinematic and acoustic markers," including syllable duration and formant profiles, were not different from those of the voluntary responses (p. 32). Chiu and colleagues (2011) interpret this result to indicate that "the production of pre-programmed gestural configurations in CV [consonant-vowel] syllables may be executed under exclusively feed-forward control." Seeking to extend their findings, these researchers (Chiu & Gick, 2014) added an additional parameter to the pre-planned syllable, asking Taiwanese Mandarin speakers to produce these syllables with different tones; the results showed that formants (and pitch contour) remained "largely unaffected by a SAS [startling auditory stimulus], supporting the view that formant profiles are pre-specified in the speech plan" (p. EL326). Unexpectedly, however, the Taiwanese participants did not correct for an upward shift in pitch related to the "physiological startle reflex, with an increase in laryngeal tension presumably acting as a protective maneuver in response to the SAS" (p. EL327; cf. Baer, 1979). Based on the assumption that "forward plans encode phonemically contrastive information," it would seem reasonable that pitch would be pre-specified for Taiwanese Mandarin, given its "phonemic role" in the language – the corollary thus seems to be that "pitch contour and pitch height are introduced at different stages of [speech] production" (p. EL327).

3.3.2 Articulatory setting theory

Comparing speech production across languages, scholars (and laypeople) have for as many as 360 years assumed that different languages may "have a whole different underlying or default posture" that could affect speech production in a foreign language (Wilson & Gick, 2014, p. 361; cf. Wallis, 1653/1972). Although this phenomenon, termed 'articulatory setting' by Honikman (1964), has received widespread theoretical attention throughout the history of phonetics (Vietor, 1884; Sweet, 1890; other terms include 'voice quality setting' and 'basis of articulation'; cf. overviews by Laver, 1978;

Jenner, 2001), it was first experimentally verified by Gick et al. in 2005 using old x-ray data. These authors found that ISPs (interspeech postures) “assumed between speech utterances: (a) are language-specific; (b) function as active targets; (c) are active during speech, corresponding with the notion of ASs [articulatory settings], and (d) exert measurable influences on speech targets, most notably including effects on the properties of neutral vowels such as schwa” (p. 231). Naturally, a phenomenon with such a widespread impact on speech articulation bears a close connection to speech motor control; prior to its experimental verification, this relationship had been hypothesized to “arise out of motor efficiency requirements” based on either “token frequency of articulatory targets” or “type frequency of articulatory targets” in a specific language (Gick et al., 2005, p. 222; Vatikiotis-Bateson & Kelso, 1990; cf. Perrier, Ostry, & Laboissière, 1996). Gick et al.’s findings support this hypothesis, and subsequent investigations have contributed more evidence for the existence of a language-specific interspeech posture (Wilson, 2006; Wilson, Horiguchi & Gick, 2007; Ramanarayanan et al., 2013; Wilson & Gick, 2014). Wilson and Gick (2014) showed that “bilingual speakers who are perceived as native in both languages exhibit distinct, language-specific ISPs, and those who are not perceived as native in one or more languages do not” (p. 361). The issue, however, “might be much more complex” than the existence of a single, specific target position, as evidenced by “significant postural differences” measured for various ISPs corresponding to different “speaking styles” in an MRI study by Ramanarayanan et al. (2013, p. 518). Articulatory setting has also often been implicated in contributing to cross-linguistic influences in Second Language Acquisition (cf. Colantoni, Steele, & Escudero, 2015), including ‘negative transfer’ and ‘fossilization’.

3.3.3 Modular accounts of speech motor control

The speech motor apparatus controlling the vocal tract needs to control a larger number of muscles than any other such functional unit in the human body (cf. the representation of various body parts on the motor homunculus). It is thus highly likely that several muscle synergies (cf. section 3.2.3 above) exist within the facial-oral-laryngeal-respiratory complex, reducing vocal tract dimensionality. While practical approaches such as midsagittal reduction and anatomical modularization have enjoyed “broad use in descriptions of speech for well over a century,” they fail to address important issues in regard to motor control, due to their unknown functional organization (Gick, 2016, pp. 177-178). Instead, Gick (2016) proposes an ‘ecological’ view that investigates the ‘dimensionality problem’ from the point of view of the organism, tying back to an early theory by Fowler et al. (1980). Improved experimental methods nowadays allow limited observation of the biomechanical properties and the neuromuscular activation of the vocal tract musculature (Bouchard et al., 2013), which can be used as a basis for modelling speech biomechanics (e.g., Nazari et al., 2011; Stavness, Lloyd, & Fels, 2012; Stavness et al., 2012; Moisik & Gick, 2013; Gick et al., in press). Results of simulations provide evidence that at least some of the speech ‘gestures’ thought to underlie speech production within the framework of articulatory phonology (Browman & Goldstein, 1986) correspond to “biomechanically optimized speech production modules” (Gick & Stavness, 2013, pp. 1-2). According to Gick and Stavness (2013), such modules organized at the neuromuscular level would be “biomechanically efficacious” (p. 1), enabling operation with little or no feedback control, often thought to be an essential requirement of speech motor control (cf. Perkell, 2012; Tourville & Guenther, 2011; section 3.3.1 above). Note that these considerations also predict the saturation effects observed in the quantal relationships observed for speech acoustics and biomechanics (section 3.3). Furthermore, a modular account of speech motor control provides principled predictions regarding the evolution and development of control structures governing the use of the vocal tract musculature, and possible transfer effects when using these muscles in different contexts (cf. section 3.5 below).

3.3.4 Individual differences in vocal tract morphology and biomechanics and their influence on speech production

No two human beings are the same and no two voices will ever sound exactly the same. A growing list of linguistic, cognitive, and sociolinguistic factors have been shown to affect an individual’s speech production (also termed idiolect), so that even monozygotic twins with almost identical vocal tract morphology and biomechanics sound slightly different. Articulatory research strives to determine the exact details of individual speech articulation and its acoustic (and auditory) consequences, and is thus often interested in precisely the subtle inter-speaker variation that is often disregarded as noise in studies concerned with principles of language that hold across groups of speakers.

This section summarizes a number of studies that have investigated the impact of individually variable morphological features on speech production. Fewer studies have looked at the biomechanics influencing speech production, defined by Perrier and Winkler (2015) as

> 1) the description of the forces or stresses acting on the body (i.e. the kinetics of the body); 2) the characterization of the intrinsic mechanical properties of the body, i.e. mass, stiffness, damping, elasticity…; [and] 3) the mathematical formulation of the physical rules determining the link between the forces and stresses applied to the body, and the time motion/deformation of the body…
> (p. 224).

Note that the authors take this definition to be more expansive than the "kinematic variables … usually measured in experimental phonetics" such as articulator "position, velocity and acceleration"; biomechanical properties of muscles are "significantly" determined by "the size and the shape of bones" through their interrelationships (anthropometry; p. 224).

One morphological feature that has received particular attention in speech production research is the shape of the hard palate. Lammert, Proctor and Narayanan (2013b) used MRI data to carry out a statistical analysis of the variability of hard palate morphology in 36 subjects (cf. Fitch & Giedd, 1999; Vorperian et al., 2005, Cheng et al., 2007); principal component analysis applied to midsagittal palate outlines yielded three main components (or modes) shown in figure 3.4 on the next page, accounting for more than 85 percent of the overall variability (p. 524). In another study from the same year (Lammert, Proctor, & Narayanan, 2013a), the same authors investigated the influence of these hard palate characteristics on the vowel productions of five speakers of American English. Acoustic measurements of the high front vowel /i/ articulated in the word "people" were compared with vocal tract simulations based on the speakers' individual morphologies "to assess the potential of different morphological variations to affect the acoustics" (p. S1926). Results showed that in the simulations, "altering the height and position of the palatal dome alter[ed] formant frequencies"; measured formant values, however, were "not significantly correlated" with palate morphology, while this was the case for "major aspects of lingual articulation" (p. S1924). This finding can be interpreted to suggest that the effects of palate shape are "not noticeable" in speech, as speakers have learned to adapt to their individual palate shape and the characteristics that would "substantially affect formant frequencies," at the same time disregarding those features "with relatively little acoustic impact" (p. S1924). More specifically, subjects appeared to "adapt their lingual contours to emulate the specific concavity and anteriority of their palates, resulting in midsagittal distance functions for all subjects that [were] relatively uniform throughout the palate region" (p. S1931). This finding provides support for acoustics as the target domain of speech motor control in vowel articulation, particularly given the fact that the authors controlled for a variety of other factors; for example, vocal tract length, which is known to "vary substantially across individuals" (p. S1924). Other work (Hiki & Itoh, 1986; Johnson, Ladefoged, & Lindau, 1993; Fuchs et al., 2006) had shown previously that mode 1 in figure 3.4, concavity, has a substantial impact on articulatory strategies (Lammert, Proctor, & Narayanan, 2013a, p. S1925).

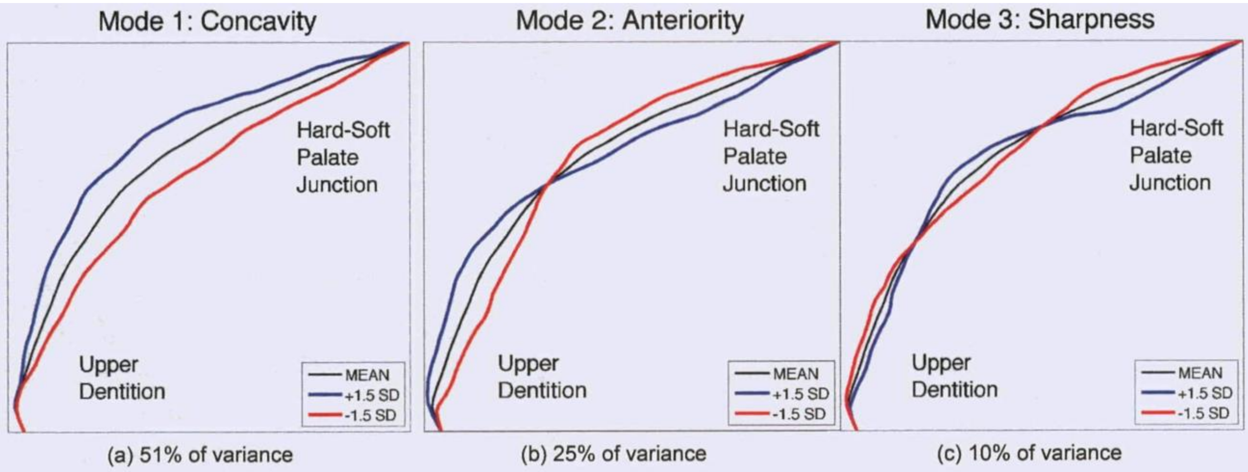


Figure 3.4: "The three largest modes of variation in hard palate shape, determined in completely data-driven fashion, without imposing any prior notions about expected shape variations, by applying PCA to the observed hard palate shapes from the subject pool. Modes reflect differences in concavity, anteriority of the apex, and sharpness of the palate around the apex. The overall mean hard palate shape is shown in black, and the blue and red lines show the nature of deviations from the mean according to each mode. The magnitude of the deviations shown reflects the magnitude of variations seen in the subject pool, at precisely ±1.5 SDs [standard deviations] from the mean shape. Because these modes account for more than 85% of the overall variance, arbitrary hard palate shapes may be well represented using only these three modes, (a) 51 % of variance; (b) 25% of variance; (c) 10% of variance." Reproduced with permission from Lammert, Proctor, & Narayanan, 2013a, p. 524.

The fact that such findings – regarding the influence of palate shape on speech production – apply across various languages was shown in an EPG study by Brunner, Fuchs, & Perrier (2009), who observed 32 speakers of five different languages "featuring different phonological characteristics" (p. 3937). The authors explored different measures of articulatory variability, with the "coefficient of variation in the percent of contact" showing a clear relationship to palate shape; "speakers with a flat palate all reduce[d] their articulatory variability in order to preserve the acoustic output," while the other speakers displayed inconsistent amounts of articulatory variability (p. 3943). In line with the findings by Lammert, Proctor and Narayanan (2013a), "acoustic variability was experimentally never found to be greater for speakers with flat palates than for speakers with domeshaped palates," adjusting "the accuracy of their tongue positioning in relation to their palate shape in order to make sure that the acoustic variability remains within a range compatible with the correct perception of the phoneme" (p. 3947).

A small number of studies have also found an effect of palate shape on consonant articulation. Rudy and Yunusova (2013) report that "[p]ositional variability of the tongue differed between the front (e.g., alveolar and post-alveolar) and back (velar) consonant groups" for 21 speakers of Canadian English, with variability of the front consonants partially explained by "palate curvature and palate length" (p. 137). Interestingly, they found no differences in positional variability related to manner of articulation, in contrast to other studies that have described fricative consonants as requiring a greater degree of articulatory precision (Alfonso & Van Lieshout, 1997; Fuchs et al., 2006; Mooshammer, Hoole, & Geumann, 2007).

Another parameter related to palate morphology is the relation of palatal length (horizontal dimension) to pharynx length (vertical dimension). This was found to influence the organization of the articulatory vowel space by Fuchs, Winkler, and Perrier (2008); see Lammert, Proctor, and Narayanan (2013b) for a statistical account of individual variation regarding pharynx shape. The comparison of two speakers with a short and long pharynx indicated that they differed in the amount of their horizontal and vertical tongue movement when producing the French low vowel /a/, such that the speaker with a short pharynx and long palate had to "control their vertical movements [more] precisely since their vocal tract constrains them to move the tongue in the vertical direction" (p. 336; cf. also earlier findings by Honda et al., 1996).

The biomechanics of speech production have only recently become the focus of experimental work and simulations, mainly because their investigation requires sophisticated imaging methods to observe and measure. One study by Perrier and Winkler (2015) used a two-dimensional biomechanical model (Perrier et al., 2003; cf. Payan & Perrier, 1997) to simulate the influence of the location of the attachment points below the temporo-mandibular joint and the resulting orientation of the styloglossus fibers on tongue biomechanics. Simulations of the tongue position for the vowel /i/ using two different models representing real speakers showed that a more vertical orientation of styloglossus fibers results in a larger vertical component in tongue position change. Comparative differences were found regarding the biomechanics of the muscles of the lips (Orbicularis Oris) on the production of rounded vowels, suggesting "that biomechanics can change the motor control strategies underlying the production of speech" (p. 248). The authors, however, warn against over-interpretation of their findings: while they are confident to have shown that "biomechanics *can* affect the elaboration of the motor control strategies, their degrees of freedom and their accuracy, it is likely" but not imperative that they also have an impact on coarticulation strategies and "*could* explain in part some trends in [speaker] idiosyncrasies" (pp. 248, 240; emphasis added).

Lastly, on the group level, there are considerable male-female differences in speech articulation. Several studies have found larger acoustic vowel spaces (on average) in females (Diehl et al., 1996; Whiteside, 2001; Simpson & Ericsdotter, 2007; Weirich & Simpson, 2014), while males have (on average) larger articulatory vowel spaces (Simpson, 2001 & 2002), although this does not affect all vowels to the same degree. This difference means that "the male tongue must traverse a greater articulatory distance at greater speed, having overall larger dorso-palatal strictures throughout the vocalic stretch" (Simpson, 2001, p. 2162) to reach acoustic targets comparable to female productions. These physiological differences have also been suggested as an explanation for the higher degrees of vowel undershoot produced by males (cf. Weirich & Simpson, 2014).

While the discussion and documentation of the impact of individual differences in vocal tract morphology and biomechanics on brass playing is beyond the scope of this thesis, I am certainly aware of the potential impact of such differences on the reported findings. If differences among brass players with different L1s, however, can be shown on the group level, this should provide a clear indication that what we are seeing is an effect unrelated to individual vocal tract morphology and biomechanics; such an effect would rather have to be explained either by language influence or the acquisition of a national style of playing.

3.4 Control of the facial-oral-laryngeal-respiratory musculature outside speech production

Speech scientists have carried out substantial work on the use of the facial-oral-laryngeal-respiratory musculature during speech production, and clinical studies have investigated its functions in basic (non-speech) behaviors “triggered by sensory stimuli (e.g., throat clearing) or associated with food intake (biting, chewing, bolus transport, sucking, clearing the oral cavity from food remainders, etc.)” (Ziegler, 2006, p. 45). Not much research, however, has been concerned with its application in voluntary non-speech activities such as whistling and more specialized abilities like trumpet or saxophone playing (Ziegler, 2006, p. 45).

According to Ziegler (2006; cf. Ziegler, 2003), there are good reasons to view these activities as separate from speech articulation, based on clinical and experimental evidence. For example, neurologic dissociations suggest “that speech motor control is organised in a task-specific way,” in agreement with ‘verticalist’ theories of the speech language interface (Liberman & Whalen, 2000) that embed speech movements into an auditory reference frame (Ziegler, 2006, p. 51). Ziegler postulates that the extensive amount of motor learning involved with speech production leads to the emergence of a “specific neural circuitry” for the control of speech articulation (p. 41), distinguishing it from other activities involving the same organs. Motor learning is commonly defined as “a set of processes associated with practice or experience leading to relatively permanent changes in the capability for skilled movement” (Schmidt & Lee, 2011, p. 327). In an earlier paper, Ziegler (2002) showed that even different speech-producing actions can differ regarding motor control; this study investigated the “potential influence of task-related factors on oral motor performance in patients with speech disorders” (p. 556). By analyzing data from 140 neurologically affected and 32 healthy control subjects, Ziegler showed that “[r]epeating a syllable at maximum rate [oral diadochokinesis] and producing the same syllable within a sentence context are two substantially different motor tasks” (p. 572). Such tasks are likely affected differentially by brain lesions, suggesting that “the concept of an underlying ‘general motor deficit’ is misleading” for the understanding of motor speech disorders (p. 573).

An opposite position to Ziegler's is taken up by McFarland and Tremblay (2006) who hypothesize that speech, swallowing and other vocal tract behaviors are "regulated through a shared network of brain regions and other neural processes that are modulated on the basis of specific task demands" (p. 300). Experimental support for their hypothesis is provided by a number of studies using neuroimaging to observe brain activation during swallowing and related motor tasks (e.g., Kern et al., 2001; Martin et al., 2004). Contrary to Ziegler's conclusion above, these authors suggest that "the oral-facial laryngeal system is organized in a manner that is largely task independent (or integrative)," as reflected by a complex system of neural control elements common to "speech, swallowing, and other orofacial movements" (p. 303). Additional support for this view is presented in the form of studies documenting the co-occurrence of speech and swallowing impairments (p. 304; cf. Müller et al., 2001; Nishio & Niimi, 2004; LaGorio, Carnaby-Mann, & Crary, 2008; Malas et al., 2015).[45]

Given that both positions presented above base their conclusions on neurophysiological observations, it is conceivable that the debate may ultimately be resolved by evidence collected at the level of biomechanics. In a recently submitted manuscript, Mayer et al. (in press) present results from a large number of simulations that indicate similarities in the neuromuscular activation shared by tongue bracing (cf. Gick et al., in press) and swallowing, which the authors interpret to imply "that aspects of the two activities must plausibly be driven by common specific sets of muscle activations" (cf. Gick & Stavness, 2013). Such an interpretation draws on a modular account of motor control (cf. section 3.2.3) and supports the proposal that speech learning might bootstrap phylogenetically encoded structures like swallowing and suckling (see Studdert-Kennedy & Goldstein, 2003; MacNeilage, 2008).

3.5 Studies of motor behavior during brass playing

Various aspects of motor behavior during brass playing have been discussed in the empirical research reviewed in section 2.5, including the recent MRI research on French Horn players by Iltis and colleagues (section 2.5.3). In this section, I provide more detail on select issues that are particularly relevant to this thesis.

[45] A case involving a pathology (tongue thrust) affecting such diverse motor behaviors as swallowing, speech production and French horn playing is documented in Fisher et al. (2014). Although the poster does not provide any information on the steps taken to alleviate the condition, it reports that the intervention succeeded in correcting the muscle patterns associated with "eating, drinking, speaking and playing [the] French horn."

In connection with the controversial tongue arch techniques outlined in section 2.4.1.2, early researchers were also interested in physical changes other than tongue position that characterize certain features of brass playing (cf. empirical research reviewed in section 2.5.1); one of non-lingual features was teeth aperture or opening as documented in studies on trumpet (Amstutz, 1970) or trombone performance (Frohrip, 1972). While Amstutz's (1970) trumpet subjects uniformly decreased teeth aperture with ascending pitch, Frohrip (1972) observed considerable variation among his trombone subjects. Due to the mandible's unlimited degrees-of-freedom, jaw movement during brass playing is potentially not limited to the vertical dimension (or the curved path similar to a trap door), although, much like speech, I assume that there is little or no movement off the midsagittal plane (cf. section 3.1.1 above). Amstutz reports that even though a majority of participants did not move their jaw horizontally, five of twenty subjects displayed this tendency in a forward direction during performance in the lower register; additionally, eight players "performed with a slight 'chewing' motion in the articulated portions of the exercises," with these movements happening rapidly and without a negative effect on performance (p. 39). Frohrip (1972) reports measurements of teeth opening (vertical position) and overbite (horizontal position) for his trombone participants. Although the author reasoned that "[t]he absence of physiological changes occurring in teeth opening and over- or underbite leads to a conclusion that such changes do not contribute to the ability of the performer to accomplish selected tasks" (p. 107), inspection of the individual measurements listed in the thesis shows that four out of nine performers decreased teeth opening, one increased teeth opening, one showed no change, and two players were variable throughout a rising pitch sequence. Calculating Euclidian distances for the combination of teeth opening and overbite yields interdental distances of 5-13.6 mm.

Only a handful of studies have placed a clear focus on motor performance during brass playing (cf. also Devroop & Chesky, 2002). For example, Bertsch and Hoole (2014) investigated tempo and endurance for tonguing on brass instruments (cf. Bertsch, 2013); a second part of the paper reports on a pilot study demonstrating the feasibility of articulographic recordings using a single subject (previously summarized in section 2.5.5 above). The main part of the study asked 206 brass players (102 amateurs, 79 brass instrument majors, and 25 professionals) to play at their maximum tempo for thirty seconds using both 'single' and 'double' tonguing (only 172 players completed this task). Single tonguing is usually accomplished by tongue motions resembling coronal stops, while double tonguing involves alternative coronal and velar contact (cf. section 2.5.3 above); players "could choose their preferred natural open note in the middle register," and "sound quality and the playing style" were not taken into account (Bertsch & Hoole, 2014, p. 412). Considerable variability was found between and within groups, and "more experienced players tend[ed] to start with tempi that they [thought] they [could] maintain for longer periods"; advanced players also "distinguish[ed] themself [sic] by more regularity" (p. 412). Overall, the "critical tempo" for single tonguing was determined to be about 120 beats per minute (BPM) when articulating sixteenth notes, although some players were able to single-tongue "up to 140 BPM and beyond" (p. 412). Overall, brass instrument majors and professionals performed at similar levels while amateurs displayed lower speed and increased variability.

Bianco et al. (2012) investigated the "temporal evolution and interrelation" of "[i]ntra-oral pressure, mouthpiece force, facial muscular activity of two groups of muscles, and the radiated sound" produced by three professional trumpet players (p. 61). The authors postulate that brass instrument performance involves a 'many-to-one' mapping (cf. Hélie et al., 1999), similar to what has been observed for speech motor control (cf. section 3.3). Accordingly, while "skillful control for each individual variable is mandatory, different input combinations are capable of delivering the same output results"; the reported results seem to support this description as "subjects performed with high reliability between repeated measurements" even though "the magnitudes of this consistency were different among the variables" (p. 61). Of particular interest is an observation regarding intensity (dynamics), namely "that two [different] motor control programs may be available for the completion of the task: a continuous feedback control for the production of lower dynamics, and a pre-planned movement control for higher dynamics," indicated by "[t]he emergence of two different shapes in the distribution of variability" (p. 62; cf. Bianco et al., 2010). The first kind of motor control program could be categorized as a closed-loop model based on the observed amount of continuous adjustments; the authors describe it as 'adaptive variability' (Kudo & Ohtsuki, 2008) typical of "skilled movements" (p. 62). In contrast, the "second control strategy (which would underlie ff [fortissimo] dynamics in isolated notes)" would produce "a fast, preprogrammed and impulsive movement" under feedforward control (cf. section 3.3.1 above), explaining "the higher variability at the sustain and final stage of the note" and "target undershoot/overshoot" (p. 62). Despite the fact that they did not observe tongue activity directly, Bianco et al. (2012) interpret "the change of concavity in the pressure profile" co-occurring with an increase in dynamics as an indication for a "shift from d-like towards t-like attacks," based on findings from speech production by Müller and Brown (1980) and Westbury and Keating (1986).

3.6 Similarities and differences between speech production and brass playing

The research reviewed above suggests several interesting parallels and differences concerning vocal tract behavior during speech production and brass playing. In this section, I will compare findings regarding the movement patterns for certain articulators, the kind of motor control underlying both activities, and explore possibilities for the transfer of movement patterns across the two activities.

3.6.1 The roles of the jaw and tongue

Concerning the role of the jaw, we know for speech production that its behavior is mostly tied to the actions of other articulators such as the tongue (often referred to as coarticulation, cf. section 3.1 above). Indeed, Perkell (1969) states that "[t]he mandible probably does not play a primary role in determining oral cavity size" (Perkell, 1969, p. 50). Keating et al. (1994) reported the following mean distances of jaw or teeth opening for vowels in English and Swedish: 5.3 mm for /i/ and 9.1 mm for /ɑ/, although /e/ is disproportionally open at 8.3 mm (p. 413). For brass playing, we know from early x-ray research that although most players seem to reduce jaw opening (or teeth aperture) with increasing pitch, patterns are highly individualistic with interdental distances ranging from 5 to 13.6 mm (Frohrip, 1972; cf. section 3.5 above). Thus, while it is possible that brass playing requires an overall lower jaw position, the findings from the data do not allow such a generalization due to the high amount of individual variability observed for both speech and brass playing. Additionally, 'chewing' (jaw movements that could be interpreted as coarticulation accompanying tongue articulation during brass playing) was documented by Amstutz for trumpet players (1970, p. 39; cf. section 3.5 above), and I have personally observed it during trombone playing, particularly in the low range of the instrument. Such motion, however, normally only happens in passages requiring fast articulation, so it should not affect sustained note production. Overall, then, it would seem that jaw position varies in similar ways during speech production and brass playing, although the more limited repertoire of articulatory motions during brass playing could mean that jaw position is less variable than during speech articulation. Usually, only one coronal place of articulation is used for brass playing (though possibly variable depending on the intended effect), and there is supposedly a single vowel-like target for sustained notes in any given register (according to pedagogical writings and previous empirical research, cf. sections 2.4 and 2.5 above). Due to tongue position being less constrained for sustained notes (based on the findings of Wolfe et al., 2010; reviewed in section 2.3 above) than for vowel articulation, the coarticulatory influence of coronal tongue movements on steady tongue position during brass playing might, however, be more considerable than the other way around (as has been shown the case for speech production, cf. section 3.1.2 above). Again, this should not affect the steady state of notes, even though the speed-accuracy trade-off seems to apply to either an articulatory (tongue position) or acoustic target (tone quality/timbre) during fast double or triple-tonguing. Bertsch and Hoole's data (2014) clearly indicates reduced tongue movement for the tongue tip sensor during fast single- and double-tonguing (figure 2 and table 1, p. 410), and regarding brass pedagogy, Zsaisits reports in his MA study on "The tongue during brass playing" that she was instructed to use /d/ and /g/ instead of /t/ and /k/ articulations during double-tonguing to reduce movement distance. Furthermore, she was told that complete constriction might not be achieved when tonguing at high speed (2012, p. 46; translation by myself). Bertsch and Hoole's (2014) data also indicate that similar biomechanical constraints might lead to tongue movement paths during brass articulation resembling the curved paths resulting from speech articulation first documented by Houde (1968; cf. section 3.3.1 above).

Showing that jaw movement patterns are similar for speech production and brass playing satisfies an important requirement for using a jaw brace to hold the ultrasound transducer in place underneath the jaw, as it ties tongue motion to jaw motion.

With regard to the tongue, it should be clear from the review of the pedagogical literature on brass playing (section 2.4) and previous empirical research on brass playing (section 2.5) that a brass player's tongue behaves similarly to vowel and coronal (and less commonly: velar) consonant articulation during speech production. Nonetheless, it should be acknowledged that this assumption relies almost exclusively on midsagittal data, and that while limited coronal data is available for speech production, this is not the case for brass playing.

The points made above regarding jaw and tongue position during brass playing and their respective variability suggest that if a default position comparable to ISP (interspeech posture; cf. section 3.3.2) exists for brass playing, it should differ from the one used for speaking. Rather, it should be optimized to achieve coronal attacks (brass players can, but rarely do, start phrases with either a breath or “ha” attack, or a velar plosive) and the respective tongue position maintained during sustained note production. If true bilinguals use separate ISPs for speaking in different languages (cf. Wilson & Gick, 2014), then a comparable ‘inter-playing posture’ (IPP) should differ from the ISP used during speech. Such an optimization should be specifically adapted to a player’s individual vocal tract morphology and biomechanics, similar to speech production, but with potential alterations accounting for higher aerodynamic pressures and the differing role of vocal tract resonances.

3.6.2 Is motor control shared across speech production and brass playing?

One criterion employed by Ziegler (2006) to distinguish speech from other motor behaviors is the degree of motor learning underlying the respective activities. While not many behaviors of the oral motor apparatus can be classified as requiring as much fine motor skill as speaking, the focus of this thesis, brass playing, would seem to satisfy this condition. There is a substantial population difference however, in that “[u]nlike whistling, singing, or saxophone playing [or brass playing], mastery of the motor skill of speaking varies only little between healthy adult individuals” (Ziegler, 2006, p. 45).

A study investigating the neurophysiological correlates of embouchure dystonia (cf. section 2.5.3 above) suggests that “an altered relationship between the hand and mouth representations in [the] somatosensory cortex seems to be related with this kind of focal dystonia” (Hirata et al., 2004, p. 817); the authors also found that players affected by embouchure dystonia “showed a decreased gap detection sensitivity of the upper lip compared with that of the lower lip” (Hirata et al., 2004, p. 818; cf. sections 2.5.2 and 2.5.6 above regarding the differing role of the upper and lower lips in brass playing). A later study by Haslinger et al. (2010) used fMRI to observe “sensorimotor activation patterns during 2 orofacial motor tasks,” comparing the performance of dystonic and unaffected players buzzing on a plastic mouthpiece, as well as in a “‘neutral’ task … simply blowing into a tube” (p. 1790). “[P]atients with embouchure dystonia showed significantly increased activation of somatotopic face representations within the bilateral primary sensorimotor cortex and of the bilateral premotor cortex” when buzzing, and “a similar activation pattern was present during the neutral task” for those players showing a clear pathology (p. 1790). These findings are in agreement with studies showing similar activation patterns in musicians with focal hand dystonia (Rosenkranz et al., 2005). Moreover, they seem to indicate that brass playing can lead to the establishment of separate cortical representations specific to this activity, possibly taking the form of task-specific muscle synergies (cf. section 3.2.3 above).

In terms of the underlying motor control models, it seems likely that both speech production and brass playing rely on closed- and open-loop (or feedforward and feedback) control in various situations. The startle research by Chiu et al. (2011) and Chiu and Gick (2014; cf. section 3.3.1) shows that single syllables can be produced using feedforward control exclusively, although it seems that pitch is not part of feedforward control in such a situation. Likewise, Bianco et al. (2012; cf. section 3.5) have shown that brass articulation at high intensities involves ballistic movements, which would suggest exclusive feedforward control relying on neurologically encoded muscle synergies and/or the application of a sliding variable (cf. sections 3.2.3 and 3.3.1, respectively).

Additionally, speech production and brass playing share an abstract or multimodal frame of reference related primarily to an acoustic output (cf. section 3.3 above), making it plausible that similar or shared mappings would exist to link articulatory gestures to sound production in both contexts. This, in turn, suggests similar challenges and solutions regarding the control of fine motor behavior involving a highly complex articulator such as the tongue.

3.7 Predictions arising from a modular theory of motor control

A meaningful comparison of the forms of motor control underlying the activities investigated in this thesis may be carried out based on a modular account of motor control, as outlined previously in sections 3.2.3 and 3.3.3. Such an approach allows the comparison of observed movement patterns, and provides principled predictions on how the hypothesized influence of native language on brass playing might arise from muscle synergies encoded in the context of speech production. Parallels concerning movement patterns have already been established in section 3.6.1 above; the rest of this section outlines a possible trajectory for the acquisition of task-specific motor control by beginning brass players.

With almost complete certainty, any brass player first starting to learn their instrument will have at this stage acquired substantial motor memory for the vocal tract movements required to produce speech in their native language; brass players usually do not start playing their instrument until they have developed their permanent incisors, although this is not an absolute requirement. Even if we disregard the challenge of developing a functioning embouchure (which may draw on muscle synergies involved in lip rounding during speech), the beginning player's vocal tract musculature faces the challenge of coming up with a way of initiating, and channeling, the required airflow into the instrument, a task "that can be executed by many possible trajectories or muscle activation patterns" (Bizzi & Cheung, 2013, p. 3). Developing a completely new sensorimotor program through a lengthy process of trial and error would be "slow and difficult" while re-using previously saved control patterns (or muscle synergies) usually provides the best strategy in such situations (Loeb, 2012, p. 761). Given the number of muscles potentially involved in this process, applying various muscle synergies from speech motor control would reduce "the volume of the space of possible motor commands that the CNS needs to search through by defining a subspace of a lower dimensionality" (Bizzi & Cheung, 2013, p. 3). Note that in our specific case, this space is also limited by a number of constraints arising from the aerodynamics and acoustics of brass playing (these will be stated precisely in the discussion section). This means that the "advantage conferred by a synergy-based control scheme" might be particularly pronounced since only a small set of motor activations might be 'good enough' to fulfill the task requirements (Bizzi & Cheung, 2013, p. 3). Furthermore, the CNS could ttest the application of "a mixture of shared and task-specific muscle synergies" (Bizzi & Cheung, 2013, p. 3); in any case, such a process would still significantly reduce the number of possible muscle combinations en route to a satisfactory solution (cf. section 3.2.4 above).

3.8 Chapter summary

This chapter began by describing the physiology of two important articulators involved in speech production and brass playing. The function of the jaw during speech production was shown to be largely co-articulatory, in contrast to the tongue, which features a large number of degrees-of-freedom and can be divided into at least two functionally independent sections. Various models of motor control were then discussed in relation to human motor behavior in general, and speech articulation in particular. This was followed by a description of the influence of individual differences in vocal tract morphology and biomechanics on speech production, and a discussion of two opposing positions concerning the question of whether there is shared motor control for speech and non-speech movements of the vocal tract. A comparison of vocal tract behavior during speech and brass instrument performance ensued, incorporating findings reviewed in previous sections. An affirmative stance was taken concerning the existence of cross-system interactions between speech production and brass playing, which provided principled predictions for language influence on brass playing based on a modular view of vocal tract motor control.

4 Possible areas of language influence on brass playing

Consideration of the issues outlined above – the physical properties underlying sound production on brass instruments (sections 2.1-2.3), brass pedagogy (section 2.4), findings of previous empirical research on brass playing (section 2.5), and accounts of motor behavior during speech production and brass playing – led me to identify two possible areas of language influence on brass playing that are described in this section.

4.1 Articulation: Beginning and connecting or ending notes

The consonants most often cited in brass pedagogy to start a note are <t> and <d>, and <k> for secondary articulations such as those occurring during double- and triple-tonguing[46]. Even though /t/ and /d/ are usually analyzed as differing only in voicing (regarding their distinctive features), their realization in different languages involves various amounts of aspiration and/or voice onset time (Lisker & Abramson, 1964; Lisker & Abramson, 1967; Klatt, 1975; Keating, Mikoś, & Ganong III, 1981; Cho & Ladefoged, 1999), and differences concerning the place of articulation or articulatory target (Kühnert et al., 1991; Lindblad & Lundqvist, 1999; Löfqvist & Gracco, 2002; Fuchs et al., 2006). For less pronounced or less 'strong' attacks, brass method books sometimes suggest using <l> or <r>; depending on a speaker's language, <l> might be realized coronally as a 'clear l' or with velarization (/ɫ/ - "dark l"), while the use of <r> in the context of brass playing most likely refers to a flap or tap (/ɾ/).

In terms of motor control, the tongue movements necessary for stop consonant production are often assumed to be under exclusive feedforward control, due to their short duration. A modular theory of speech motor control hence predicts that such movements would rely heavily on neurologically encoded muscle synergies (cf. Gick & Stavness, 2013), making it highly likely that such muscle synergies would be re-used in the context of brass playing, where similar movements are used to begin and connect/end notes. If such transfer of motor actions indeed occurs from speech production to brass playing, subtle differences in the realization of similar consonants in two languages should lead to distinctions in the transients produced at the beginning and during the connection or at the end of notes (cf. sections 2.3.2 and 2.6; Diaz,

[46] I am using <x> for letters/spelling here to indicate that this is the level of reference commonly used in brass pedagogy texts, and to distinguish it from phonological/phonetic descriptions.

2011) when executed by players with different L1s. Such acoustic differences should be perceivable at least by expert listeners (cf. section 2.6).

4.2 Tongue position during sustained note production

Since at least 1584 (Dalla Casa, 1584/1970), teachers of brass instruments have advocated the use of vowel tongue positions for brass playing. Based on the conviction that there is an effect of vocal tract resonances on brass instrument sound (cf. section 2.3), tongue position and corresponding changes in the pharyngeal cavity should affect the timbre of any note produced by brass players. The tongue movements associated with vowel production usually exhibit longer durations than for stop consonant production, and incorporate auditory (as shown by acoustic perturbation studies, e.g. Purcell & Munhall, 2006a&b; Villacorta, Perkell, & Guenther, 2007) and (likely) sensory feedback. Nonetheless, the muscular complexity of the tongue (cf. section 3.1.3) would seem to suggest that the motor system makes use of a strategy similar to the one outlined in section 3.7 above when searching for an adequate tongue position to be used during sustained note production.

Thus, if transfer effects happen from speech production to brass playing, a varying number of more or less closely spaced vowel tongue positions should be available to speakers of different native languages when playing their instruments, encoded in the form of muscle memory and/or pre-programmed muscle synergies. The tongue positions assumed during playing need not be completely identical to a vowel tongue position but could be limited to a certain functional area of the tongue (cf. section 3.1.3 above) or movement patterns encoded by muscle synergies, while other sections of the tongue might be more closely constrained by the aero-dynamical or acoustic demands of brass playing. The utilization of vowel tongue positions affecting the whole tongue shape or only a specific functional section of the tongue should lead to a characteristic timbre during sustained note production on brass instruments.

4.3 Hypotheses

Hypothesis 1: Brass players can perceive (consciously or subconsciously) the acoustic consequences of playing differences between players with different native languages.

Hypothesis 2: Tongue positions assumed during sustained note production on brass instruments are based on motor memory.

a) This motor memory will be based on speech articulation, specifically the tongue shape of vowels.

b) Functionally independent sections of the tongue will be individually affected by motor memory from a player's native language, in agreement with a modular theory of motor control.

These hypotheses will be addressed with an online questionnaire whose findings are presented in the following chapter, and an ultrasound study of Tongan and New Zealand trombone players described in the remainder of this thesis.

5 Preliminary investigation: Online questionnaire

To determine whether a majority of brass players report that they can perceive differences in the sound produced by players with different native languages (*hypothesis 1* above), I conducted an online questionnaire that asked questions such as “Do you think that a person’s first language and/or other acquired languages have some kind of influence upon playing brass instruments?,” as well as collecting data on respondents’ musical back ground and playing proficiency (cf. Heyne & Derrick, 2015d). Over 300 brass players from all over the world answered the questionnaire comprised of 25 questions (taking about roughly 15-20 minutes to complete) and available in both English and German; a high number of incomplete responses, however, mean that the statistical analysis below is based on a subset of 135 complete observations only. Questionnaire participants included in the statistical models were of 24 different nationalities and spoke 15 different native languages (including some bilinguals), although the majority of responses (107) came from speakers of either English or German. The mean age was 36 years and the sample seems to be representative of the gender imbalance in brass playing with 96 males versus only 39 females submitting complete responses. In terms of playing proficiency, roughly half of the participants indicated they were professional players, with a quarter each reporting amateur and semi-professional abilities; the mean number of years playing a brass instrument was quite high, roughly 22 years, and respondents reported to play their instruments for a mean duration of almost 16.5 hours per week. Table 5.1 below provides more detail on participant demographics, as well as the exact number of responses to the two central questions regarding language influence on brass playing: (1) Do you think that a person’s first language and/or other acquired languages have some kind of influence upon playing brass instruments?, and (2) “Which factor do you think is more influential in affecting brass playing, one’s First Language/s (and possibly Second Languages) or playing styles (nationals schools etc.)?” See appendix A, pages 236-242 for the full questionnaire.

complete responses	135	**Do you believe that national schools/styles exist regarding (classical) music in general or specific instruments/groups of instruments?**	
different nationalities	24	yes	68
different native languages (includes 4 bilinguals)	15	probably	37
different varieties of English (counted as one language)	6	undecided	26
males	96	probably not	3
females	39	no	1
mean age	36.05		
		(1) Do you think that a person's first language and/or other acquired languages have some kind of influence upon playing brass instruments?	
proficiency:		no influence whatsoever	18
professional	64	limited influence	37
semi-professional	38	influences playing somewhat	46
amateur	33	clear influence	23
		much influence	11
		(2) Which factor do you think is more influential in affecting brass playing, one's First Language/s (and possibly Second Languages) or playing styles (nationals schools etc.)?	
mean number of years played	21.88	Style has most influence	48
mean playing time per week	16.46	Style is somewhat more influential	40
		undecided	34
mean number of L2s spoken	1.59	Language is somewhat more influential	12
		Language has most influence	1

Table 5.1: Selected demographic and response data from the online questionnaire.

To determine which factors might influence a respondent's answer to the two central questions mentioned above, I fit separate cumulative link models (clm: Christensen, 2015) in a stepwise backwards iterative fashion in R (R Development team, 2015) until all remaining predictors reached significance; only main effects were considered and all non-numeric predictors where treated as nominal factors. These models were back-fit along the Akaike information criterion (AIC) to measure quality of fit; this technique allows for the isolation of a statistical model that provides the best fit for the data. In a first step, all possible predictors were tested for covariance, resulting in the removal of age (correlated with the number of years played), playing proficiency, and source of income (involving brass instrument performance, teaching, or something unrelated; both proficiency and soure of income were correlated with weekly playing time). Other factors had to be removed completely (style of music most frequently played; country where brass playing was learned; dominant style of playing learned: brass band, jazz, etc.; L1; nationality; and sex, due to a limited age range for the females) or reduced in their levels (trombone was added to 'low brass' for instrument played; no, problably not, and undecided were conflated into one category for the answer to the following question: "Do you believe that national schools/styles exist regarding (classical) music in general or specific instruments/groups of instruments?") due to a lack of data and/or an insufficient balance of observations across categories.

To arrive at the final models shown in tables 5.2 and 5.3 below, a number of predictors were additionally removed as they turned out to be non-significant. For the model on language influence these were: number of years played, number of L2s spoken, whether respondents were told to use speech sound configurations in their lessons, the answer to question (2) regarding the influence of language versus that of style, and the instrument played. In the case of the model on the perceived importance of language versus style all predictors except the possible answers to question (1) regarding the influence of language influence on brass playing were removed as they were non-significant and did not add to the explanatory power of the model as determined by ANOVA comparisons using the Akaike criterion.

The first model informs us that brass players who spend more time playing their instruments (highly significant) are more likely to believe in the influence of native language on brass playing. Depending on which national school of playing (including 'not reported' for participants who did not provide this kind of information) was set as default for the category 'schools learned,' various national schools or styles of playing also came out as significant or almost significant – see figure 5.1 and table 5.3 for details. Furthermore, there is weak correlation with the belief in the existence of national schools/styles of playing. Recall that the answers to the question "Do you believe that national schools/styles exist regarding (classical) music in general or specific instruments/groups of instruments?" were conflated into three categories (see above), with 'probably' serving as the default.

The second model indicates that respondents are more likely to rate the influence of style as more important than that of language if they answered question (1) negatively ("no influence whatsoever").

Formula: language_influence ~ playing_time + schools_learned + national_schools					
	Estimate	SE	Z value	Pr(>|z|)	Signify-cance
Coefficients:					
weekly playing time	0.05924	0.01725	3.433	0.000596	***
schools learned: American	0.25963	0.57469	0.452	0.651439	
schools learned: Aus-NZ	-0.51699	0.87058	-0.594	0.552611	
schools learned: Austrian	-1.51461	1.64431	-0.921	0.356987	
schools learned: British	-0.73814	0.65394	-1.129	0.259000	
schools learned: French	0.64957	1.02105	0.636	0.524661	
schools learned: German	-1.17365	0.59623	-1.968	0.049014	*
schools learned: Italian	-0.58532	1.67179	-0.350	0.726254	
schools learned: Jazz	0.84035	0.76220	1.103	0.270228	
schools learned: multiple	-0.86505	0.54594	-1.585	0.113079	
schools learned: Scandinavian	-1.52357	1.25535	-1.214	0.224876	
existence of national schools of playing: undecided/probably not/no	-0.70439	0.47742	-1.475	0.140100	
existence of national schools of playing: yes	0.69112	0.40853	1.692	0.090698	.
Threshold coefficients:					
no influence whatsoever | limited influence	-1.5208	0.5869	-2.591	N/A	N/A
limited influence | influences playing somewhat	0.2687	0.5638	0.477	N/A	N/A
influences playing somewhat | clear influence	2.1149	0.5963	3.546	N/A	N/A
clear influence | much influence	3.6659	0.6594	5.559	N/A	N/A
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 5.2: Model outputs for the cumulative link model fit in R to predict the response to the question (1) Do you think that a person's first language and/or other acquired languages have some kind of influence upon playing brass instruments?

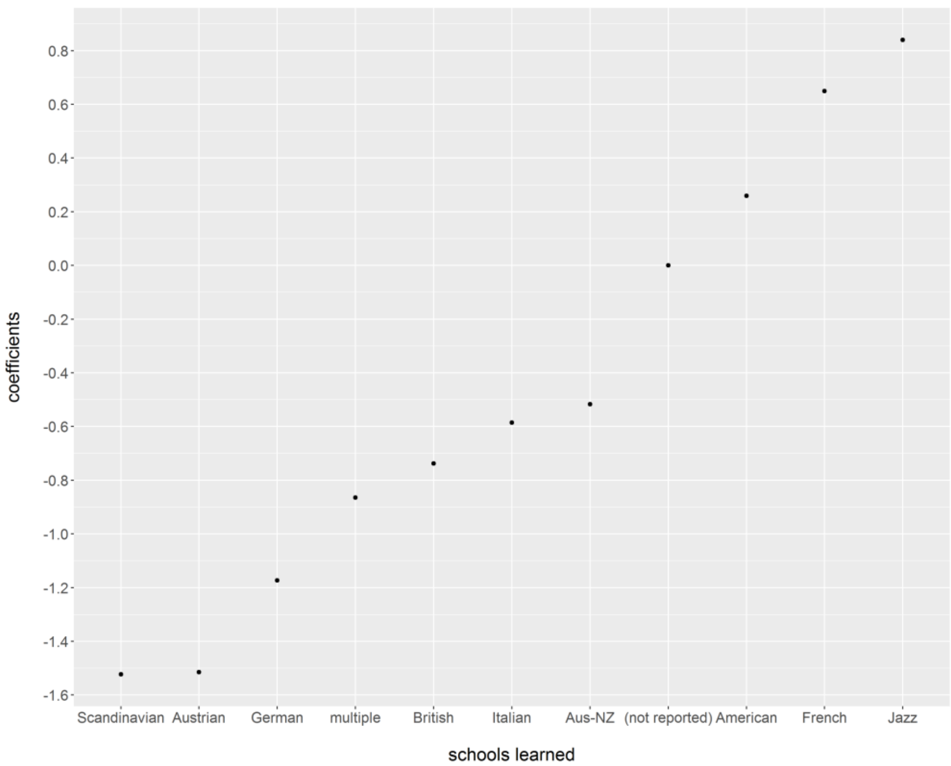


Figure 5.1: Ordered plot of the coefficient values for the various categories of ‘schools learned’ as estimated by the best fit-model on language influence reported in table 5.2 above.

		schools learned										
schools learned: default	(based on reported model)	(not reported)	Scandinavian	Austrian	German	multiple	British	Italian	Aus-NZ	American	French	Jazz
(not reported)	N/A	N/A	0.25	0.36	0.05 *	0.11	0.26	0.73	0.55	0.65	0.52	0.27
Scandinavian	-1.524	0.22	N/A	1	0.78	0.59	0.54	0.64	0.47	0.15	0.15	0.08 .
Austrian	-1.515	0.36	1	N/A	0.83	0.69	0.64	0.68	0.57	0.28	0.23	0.17
German	-1.174	0.05 *	0.78	0.83	N/A	0.56	0.51	0.72	0.45	0.01 *	0.07 .	0.01 *
multiple	-0.865	0.11	0.59	0.69	0.56	N/A	0.84	0.86	0.68	0.02 *	0.13	0.02 *
British	-0.738	0.26	0.54	0.64	0.51	0.84	N/A	0.93	0.81	0.13	0.19	0.06 .
Italian	-0.585	0.73	0.64	0.68	0.72	0.86	0.93	N/A	0.97	0.61	0.5	0.41
Aus-NZ	-0.517	0.55	0.47	0.57	0.45	0.68	0.81	0.97	N/A	0.36	0.34	0.18
American	0.26	0.65	0.15	0.28	0.01 *	0.02 *	0.13	0.61	0.36	N/A	0.7	0.44
French	0.65	0.52	0.15	0.23	0.07 .	0.13	0.19	0.5	0.34	0.7	N/A	0.87
Jazz	0.84	0.27	0.08 .	0.17	0.01 *	0.02 *	0.06 .	0.41	0.18	0.44	0.87	N/A
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1												

Table 5.3: Significance table for the various categories of ‘schools learned,’ determined by re-levelling the various factors.

Formula: lg_vs_style ~ language_influence					
	Estimate	SE	Z value	Pr(>|z|)	Signify -cance
Coefficients:					
language influence: influences playing somewhat	-0.3725	0.4698	-0.793	0.42781	
language influence: limited influence	-0.7324	0.4896	-1.496	0.13466	
language influence: much influence	-0.4915	0.6924	-0.710	0.47776	
language influence: no influence whatsoever	-2.0327	0.6633	-3.065	0.00218	**
Threshold coefficients:					
Style has most influence | Style is somewhat more influential	-1.23738	0.40862	-3.028	N/A	N/A
Style is somewhat more influential | undecided	0.05385	0.39994	0.135	N/A	N/A
undecided | Language is somewhat more influential	1.69929	0.45254	3.755	N/A	N/A
Language is somewhat more influential | Language has most influence	4.38910	1.05555	4.158	N/A	N/A
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 5.4: Model outputs for the cumulative link model fit in R to predict the response to the question (2) "Which factor do you think is more influential in affecting brass playing, one's First Language/s (and possibly Second Languages) or playing styles (nationals schools etc.)?"

Overall, the questionnaire results show that brass players believe they can perceive differences in the playing of colleagues with different native languages. Whether this means that they would be able to tell players from different groups apart based only on the produced sound is of course another question (cf. findings by Cox reported in section 2.7). Hypothesis 1 "Brass players can perceive (consciously or subconsciously) the acoustic consequences of playing differences between players with different native languages" can thus be answered cautiously in the affirmative for the purposes of this study, although future perceptual research is clearly needed to properly address this question.

6 Methodology: Ultrasound imaging of the tongue (UTI)[47]

Although the technique was already mentioned in section 2.5.4 above, this section will provide a brief overview of the physics underlying ultrasound imaging and its application to speech research. Ultrasound for medical imaging uses ultra-high frequency sound ranging from 3-16 MHz to penetrate soft tissues and calculate an image of their density by evaluating the echo returned when sound waves get reflected due to changes in tissue density; it was first applied to speech research by Sonies et al. (1981). To produce ultrasound signals, ultrasound machines use piezoelectric crystals embedded in a transducer (or probe), which is held underneath the chin when performing lingual ultrasound. Ultrasound waves "get absorbed by bone and reflect sharply off of air boundaries," meaning that the technique does not image bone or air very well (Gick, Wilson, & Derrick, 2013, p. 161); its second property, however, is very useful for speech research as it provides good resolution of the tongue surface as long as there is continuous tissue for the sound waves to travel through. The most common application of ultrasound in speech research is along the midsagittal plane, which in theory should provide a continuous image of the tongue surface bounded by the hyoid bone towards the back of the tongue and the mandible at the front; nonetheless, the tongue tip is almost never completely visible in ultrasound images due to an air pocket that forms underneath it when the front of the tongue is lifted from the floor of the mouth. Another option is to turn the ultrasound probe 90 degrees and record coronal images for a given slice of the tongue. Modern ultrasound machines can provide a temporal resolution of up to 120 Hz for two-dimensional images, although it remains problematic to actually capture data at this rate as ultrasound machines lack the capacity to record sufficient lengths of videos internally and with synchronized audio. The above described properties of the technique in terms of tissue resolution furthermore mean that hard structures in the oral cavity, such as the teeth or the palate, do not show up on ultrasound images, making it an absolute necessity to stabilize the ultrasound transducer in relation to the head if images are to be compared across time and/or subjects. Notwithstanding its drawbacks, ultrasound imaging provides a noninvasive and relatively inexpensive method for imaging the tongue when compared to alternatives like x-ray imaging, electromagnetic articulography and MRI (cf. section

[47] This section builds on information from Gick, Wilson, & Derrick (2013) and Stone (2005); the latter reference also provides a comprehensive in-depth discussion of all the issues to consider when using ultrasound imaging to measure tongue position and motion.

2.5 above). Recently, systems enabling three-dimensional imaging have become available and have seen their first applications to speech (Lulich, 2014; Lulich, 2016) and clarinet performance (Lulich, Charles, & Lulich, 2016).

7 Data collection

Hypothesis 2 outlined in chapter 4 above predicted that tongue positions assumed during sustained note production on brass instruments are based on motor memory and that furthermore, a) such motor memory would be based on speech articulation, specifically the tongue shape of vowels, and b) functionally independent sections of the tongue would be individually affected by motor memory from a player's native language, in agreement with a modular theory of motor control. To test these assumptions, I decided to record participants' tongue positions and movements while speaking in their native language and playing the trombone. This chapter describes how and why wordlists and particular musical passages were chosen to elicit these behaviors, how participants were selected, and the measures that were taken to limit the influence of various factors that can influence brass instrument sound and speech production.

7.1 Selection of instrument group to record

Even though I initially believed that the findings from this study would apply to all brass instruments, realizing that a multitude of factors can affect brass instrument sound (even if one disregards vocal tract influence, discussed in section 2.3 above) meant that I would have to try to control for as many variables as possible. One consequence of this was to limit the study to a single instrument, and naturally the instrument of choice was my own primary instrument, the trombone. Trombonists these days can choose to perform on different instruments produced by a large number of manufacturers, built of various materials and with varying physical dimensions, both of which influence the sound produced by the instrument. Parameters that have been shown to influence the sound spectrum of brass instruments are material (Smith, 1986; Pyle Jr., 1998), taper (or conicity) of the bore (Ayers, Eliason, & Mahgerefteh, 1985), the flaring shape of the bell (Campbell & Greated, 1994; Campbell, Myers, & Chick 2013), and the shape of the mouthpiece (Caussé, Kergomard, & Lurton, 1984; Wright & Campbell, 1998; Carral & Campbell, 2002), which interacts with other aspects of instrument design (Benade, 1976, p. 401; Fletcher & Rossing, 1991, pp. 371-372, 375).

Naturally performers have different personal preferences regarding the characteristics of the instrument they choose to play, guaranteeing that any particular instrument chosen for this study would be met with approval by some participants, while being deemed unacceptable by others. Luckily, in 2011, the world's first plastic trombone ('pBone') was released by Warwick Music, Ltd. in the UK, and I decided to use this instrument for my research as it was unlikely to be met with disapproval due to its novelty value. Furthermore, many of the participants would not have played such an instrument before or at least would not play a plastic trombone regularly, reducing the impact of familiarity with the test instrument on my experimental findings. Nonetheless, a potential confound remained with some participants normally performing on small-bore tenor trombones (somewhat similar in feel to the 'pBone'), with others more commonly playing on large-bore orchestral tenor trombones, or bass trombones.

7.2 Selection of participant languages

NZE as the local variety of English spoken in Christchurch was an obvious choice as one of the languages to be included in this study, but not just for practical reasons. Before recording the first participants for this study, I thought that a central vowel tongue position (schwa /ə/) would be an obvious candidate for a tongue shape to be used during trombone playing, due to its frequent descriptions as being relatively un-constricted (but see section 10.4.1 for findings challenging this notion). NZE features a second centralized vowel position in addition to schwa, with the KIT vowel having centralized as part of the characteristic short front vowel shift (see section 7.3.1 below). Its inclusion also made sense based on NZ's historically close ties to Britain and the establishment of the British orchestral and brass band traditions in the former colony; the current local playing style of brass playing can be understood to be British in essence, although global developments (cf. section 2.4.2 above) have probably rendered it more 'international' in the last few decades.

The second language chosen for this study is Tongan, a Malayo-Polynesian language spoken by 126,390 people (Kuo & Vicenik, 2012; drawing on Lewis, 2009) in the Kingdom of Tonga, situated in the South Pacific. Tongan phonology differs markedly from (NZ) English in terms of its vowel system and consonant articulation, as well as supra-segmental phonology. The underlying phonological unit in Tongan is the mora and prominence is indicated by a pitch accent, lower F1, longer duration and higher energy for vowels with primary stress, as well as a pitch accent plus shorter duration for vowels with secondary stress (Garellek & White, 2015, p. 30). Tonga has a huge British-style brass band tradition and at the time of Aldred's Master's thesis (1997) on *Ifi palasa: Brass bands in Tonga* (Ifi palasa means 'brass blowing' in Tongan), there were “more brass bands in Tonga, per head of population, than in the North of England where the brass band movement began. Tonga has possibly the highest density of brass bands in the world” (Aldred, 1997, p. 12).

7.3 Brief overview of NZE and Tongan phonetics and phonology

7.3.1 NZE

NZE is the local variety of English spoken in NZ, and the English language was introduced to NZ by settlers from Britain arriving in large numbers from about the middle of the 19th century. Modern NZE is non-rhotic, reflecting the high percentage of original settlers who originated from the South-East of England (Hay, Maclagan, & Gordon, 2008). It is the only variety of English whose pronunciation can be traced back all the way back to its beginnings, as documented through the Origins of New Zealand English project; the project analyzed early recordings made in 1946-48 of the first generation of New Zealanders born in the country (Gordon, Maclagan, & Hay, 2007). Foreigners often have trouble distinguishing NZE from another Southern hemisphere variety, Australian English (AusE), although they differ markedly in their short front vowel realizations (Hay, Maclagan, & Gordon, 2008, p. 14). AusE has very close KIT vowel (Cox & Palethorpe, 2008) while in NZE, this vowel has centralized as part of a characteristic chain shift, outlined in figure 7.1 on the next page (from Hay, Maclagan, & Gordon, 2008, p. 41)[48]. This chain shift was allegedly initiated by a relatively high TRAP vowel brought to NZ by the first European settlers; this vowel continued to rise, encroaching on “the acoustic space of DRESS, which in turn rose, crowding KIT” (Hay, Maclagan, & Gordon, 2008, p. 42). While it would have been possible for KIT to become more close as in AusE, in NZE, it instead centralized and subsequently lowered, assuming its present-day location (Hay, Maclagan, & Gordon, 2008, p. 42).

[48] Throughout this thesis, the lexical sets developed by Wells (1982) will sometimes be used to refer to the vowel qualities of NZE. For a mapping of these keywords to the corresponding IPA symbols as used for NZE in this thesis see figure 7.2 below.

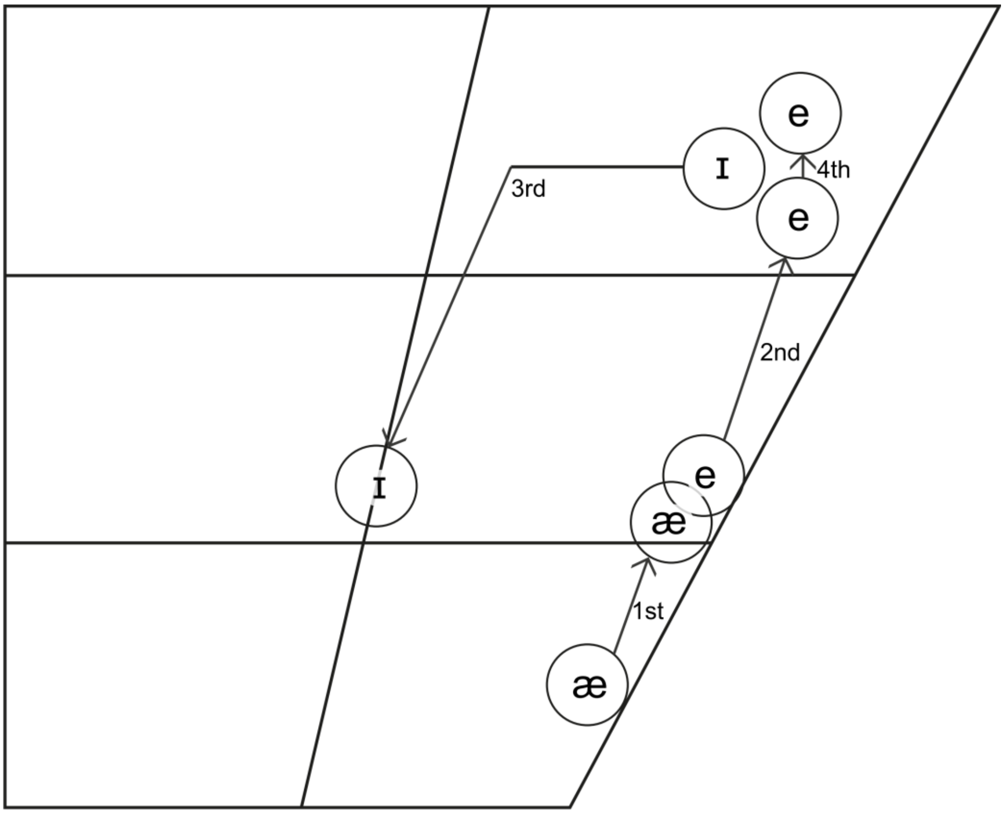


Figure 7.1: The New Zealand English short front vowel shift. Note that this figure is reprinted here as a mirror image to match the conventional orientation of the articulatory data presented in the results section.
Hay, J., Maclagan, M., & Gordon, E. (2008). *New Zealand English*. Edinburgh: Edinburgh University Press; p. 41, reproduced with permission of Edinburgh University Press via PLSclear.

Ongoing movement of the DRESS vowel suggests that this vowel shift is still in progress; with DRESS having risen even further, “DRESS and FLEECE now completely overlap in acoustic space for many young speakers, and for some innovative individuals DRESS has risen above FLEECE and can be more front than FLEECE” (Maclagan & Hay, 2004, p. 187). As a consequence, FLEECE is becoming increasingly diphthongal, with innovative speakers producing “pronounced on-glides” to maintain a phonemic distinction of the two vowel phonemes, implicating FLEECE in the New Zealand ‘short front vowel’ shift (Maclagan & Hay, 2004, pp. 188, 183). Figure 7.2 on the following page shows the “a generalisation of roughly where” the monophthongs of NZE “are produced by New Zealanders,” although the authors acknowledge considerable variation “across speakers and context” (Hay, Maclagan, & Gordon, 2007, p. 21). The image does not include unstressed schwa, although its location can be inferred to be very close to KIT, based on the following comment from Hay, Maclagan and Gordon (2008): "One effect of the centralisation of KIT in NZE is that there is almost no audible difference between KIT and the neutral vowel schwa, usually written /ə/" (p. 23); similarly, Bauer and Warren (2004) state that "there is no phonetic distinction between the KIT vowel and commA vowel where that occurs in non-final position," and they suggest the use of the same phonetic symbol for both sounds (/ə/; p. 587).

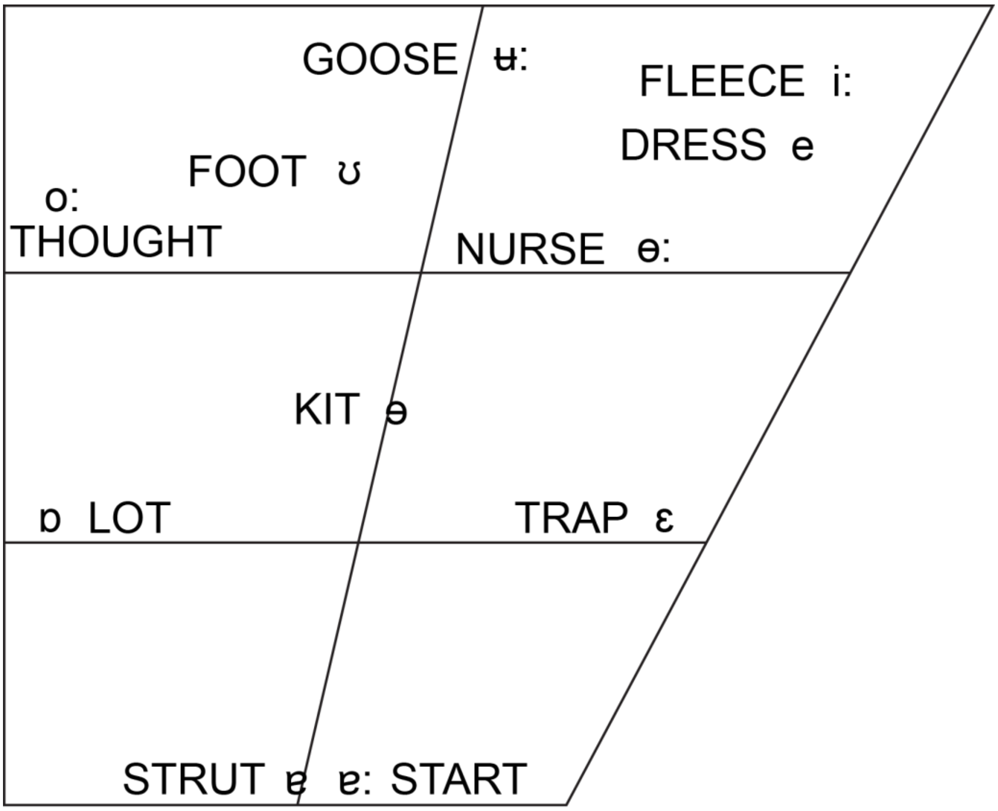


Figure 7.2: The monophthongs of New Zealand English. Note that this figure is reprinted here as a mirror image to match the conventional orientation of the articulatory data presented in the results chapter.

Hay, J., Maclagan, M., & Gordon, E. (2008). *New Zealand English*. Edinburgh: Edinburgh University Press; p. 21, reproduced with permission of Edinburgh University Press via PLSclear.

7.3.2 Tongan

	Bilabial	Labio-dental	Dental	Velar	Glottal
Plosive	p		t	k	ʔ
Fricative		f v	s		h
Nasal	m		n	ŋ	
Lateral approximant			l		

	Front	Central	Back
Close	i		u
Mid	e		o
Open		a	

Table 7.1: Tongan consonant inventory (top) and vowel inventory (bottom).
Garellek, M., & White, J. (2015). Phonetics of Tongan stress. *Journal of the International Phonetic Association*, 45(1), p. 15, reproduced with permission.

The phonological system of Tongan consists of twelve consonant and five vowel phonemes (Garellek & White, 2015, p. 15), illustrated in Table 7.1 above. Only the labiodental fricatives /f/ and /v/ are distinguished by voicing, and the coronal consonant series is dental; vowel length is contrastive, and "[s]tress … falls on the penultimate mora, where single vowels consist of one mora and so-called long vowels (which are perhaps best considered to be two single vowels in a row) consist of two morae" (Kuo & Vicenik, 2012, p. 64). Tongan also features an unusual phenomenon referred to as 'definitive accent', which has been analyzed "in various ways in the literature: as stress shift from penultimate to final vowel, as simultaneous stress reduction on a penult and stress addition on an ultima, and as addition of a syllable by repetition of the final vowel" (Anderson & Otsuka, 2006; cf. Churchward, 1953; Taumoefolau, 2002). Furthermore, Garellek and White (2015) found that vowels with primary, secondary or no stress differ in their first formant; see figure 7.3 on the next page. Both observations, however, are not directly relevant to this study, as stimuli were presented as wordlists (see section 7.4.2 below), and only the articulation of primarily stressed (or accented) vowels was measured.

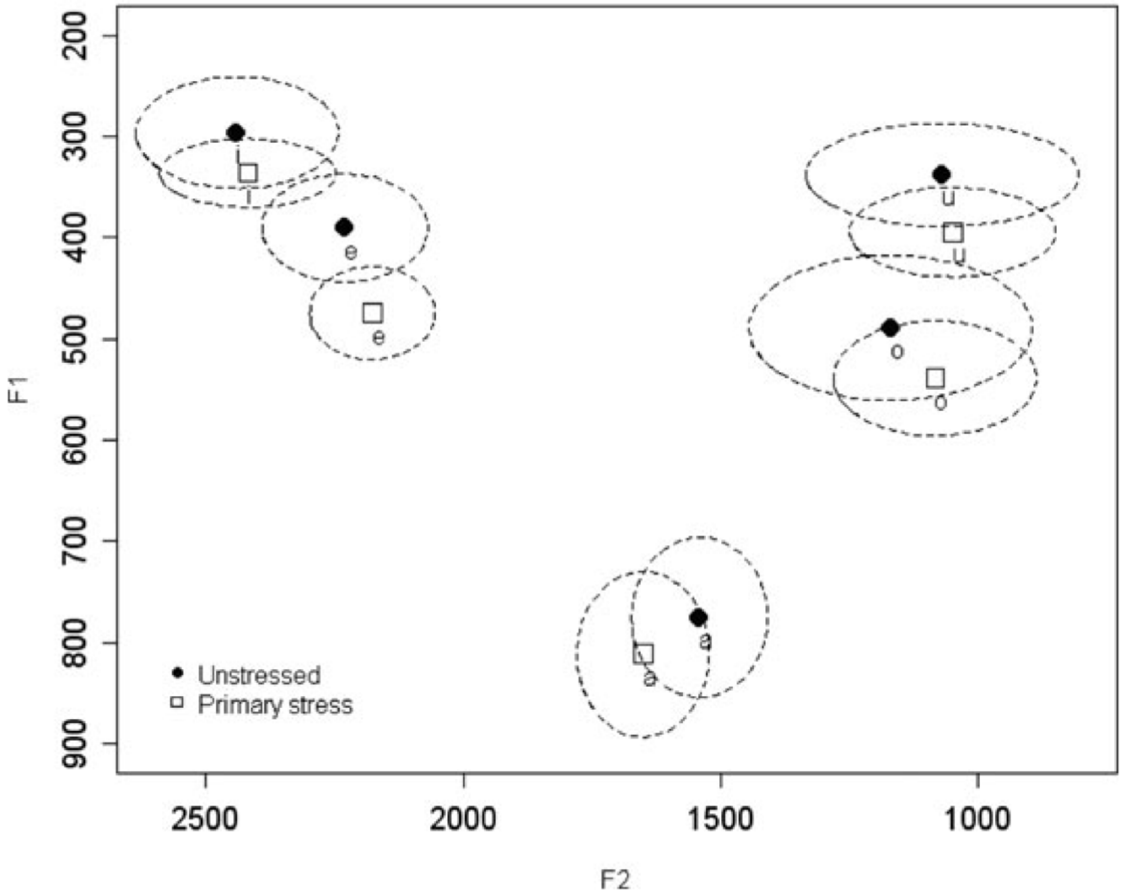


Figure 7.3: Vowel plots for primary stress versus unstressed vowels. F1 x F2 clouds show one standard deviation from mean value.
Garellek, M., & White, J. (2015). Phonetics of Tongan stress. *Journal of the International Phonetic Association*, 45(1), p. 22, reproduced with permission.

7.4 Speech elicitation: Wordlists

7.4.1 NZE wordlist

Except for the pilot participant, all NZE-speaking participants were asked to read a list of 803 words off a computer screen (participant S3 read an earlier, slightly longer version of the final list). Words were presented in blocks of three to five items using Microsoft Powerpoint, with the next slide appearing after a pre-specified, regular interval; this time-step was inserted to elicit possible interspeech postures between blocks after I developed the hypothesis that a possible 'inter-playing posture' assumed during rests in brass playing might be related to interspeech postures occurring during speech pauses (cf. sections 3.3.2 and 3.6.1 above). Note that the revised elicitation paradigm did not necessarily lead to a pause between blocks, but depended on each participant's reading speed and the number of words presented on a given slide.

Additionally, the syllables /tatatatata/ (or /dadadadada/) were displayed at the beginning and end of each block to elicit coronal productions that would allow me to correct for audio-video misalignment by temporally aligning tongue movement with the resulting rise in the audio waveform (Miller & Finch, 2011, cf. section 8.2 below).

Words were chosen to elicit all eleven monophthongs of NZE in stressed position plus unstressed schwa, and occurring in combination with preceding coronal (/t, d, n/) and velar (/k, g/) consonants, as well as rhotics and laterals. Although it is well-known that read speech and wordlists result in somewhat unnatural speech production (Barry & Andreeva, 2001; Zimmerer, 2009; Wagner, Trouvain, & Zimmerer, 2015), this form of elicitation was chosen to ensure that the desired phoneme combinations were reliably produced, and to facilitate the automatic segmentation of the recorded speech data based on acoustics. While the blocks usually contained words with the same stressed consonant-vowel combination, the sequence of the blocks was randomized so participants would not be able to predict the initial sound of the first word on the following slide. For all but the pilot participant, this procedure resulted in nine blocks of speech recordings lasting roughly two minutes and twenty seconds each.

The pilot participant read a list of words of similar length printed on paper and presented in lines of three to seven items; unfortunately, much of his data had to be discarded due to movement of the ultrasound transducer (see section 8.7.1).

7.4.2 Tongan wordlist

The same setup for token presentation was used for the Tongan speakers; however, assembling an appropriate wordlist was more challenging due to my limited expertise in Tongan. Fortunately, Tongan spelling is mostly phonetic, enabling me to use Tu'inukuafe's simplified dictionary of modern Tongan (1992) to compile a list of 1,154 words to elicit all five vowels of Tongan, both as short and long vowels, and occurring in combination with all relevant coronal and velar consonants. Even though this process let to a selection of a high number of English loanwords, I do not believe that it affected the comparability of the experimental materials; in any case the articulatory results for the two language groups reported below should have become more similar, reducing the findings of the study.

Note that Tongan words can be quite short, consisting minimally of a single vowel phoneme, so that it did not take as long to elicit the Tongan wordlist as the numerically shorter NZE wordlist. Copies of both wordlists are included in appendix B, pages 244-250.

7.5 Musical passages

The musical passages recorded for this thesis were designed to elicit a large number of sustained productions of different notes within the most commonly used registers of the trombone. Notes were elicited at different dynamics and with various articulations including double-tonguing, which features a back-and-forth motion of the tongue to produce coronal and velar articulations (cf. ‘lingua reversa’ in section 2.4.1.1; section 3.6.1). To control as much as possible for the intonation of the produced notes, five out of a total of seven passages required only notes played in first position, and participants were asked to ‘lock’ the slide for this part of the recordings (the slide lock on a trombone prevents extension of the slide); for the same reason, the tuning slide of the ‘pBone’ was completely pushed in for each recording session, and participants were required to perform using the exact same mouthpiece, a standard 6 1/2 AL mouthpiece by Arnold’s and Son’s, Wiesbaden, Germany. Note that the pilot participant performed on his own ‘pBone’ (there is only one model), using a personal mouthpiece with a bigger size than the one supplied to all subsequent participants.
The difficulty of the selected musical passages was quite low to ensure that even amateurs could execute them without prior practice. The first six exercises consisted only of held notes and lip slurs, such as those often used for warming up on the instrument, but restricted to the harmonic series playable in the first slide position. The final two exercises required limited use of the slide, and were slightly modified from original etudes written for brass instruments (see appendix B, pages 244-250).

7.6 Recording locations

All participants recorded in Christchurch were asked to perform in a small sound-attenuated room on campus, chosen to limit the effect of room acoustics on the recorded sound (cf. Kuttruff 2014). Unfortunately, no equivalent room was available on the campus of the Royal Tongan Police Band in Nuku’alofa, capital city of the Kingdom of Tonga. The room used in Tonga was roughly one-and-a-half times the size of the room at the University of Canterbury and did not feature sound attenuation; additionally, the sound from a nearby workshop, a gym, and chickens living on campus sometimes made its way into the recordings, reducing the potential of acoustic comparisons of recordings made in the two locations. Although there exists a small confound in the fact that the room acoustics may have affected the way participants played the selected musical passages, no perceptual research was carried out using the recordings as stimuli, which would have been problematic. The room situation in Tonga also meant there was no chance to control for ambient temperature, which has been shown to influence the speed of sound and thus, wind instrument acoustics (Gilbert, Ruiz, & Gougeon 2006). Temperature differences between rooms, however, were not too pronounced due to it being winter in Tonga, and were thus probably negligible compared to changes in $C0^2$ concentration and temperature within the instrument bore, as caused by the composition of the players' breath (cf. section 2.2 above).

7.7 The participants

A total of 29 trombone players with six different native languages were recorded for this thesis; however, not all of the data could be analyzed and only participants from the two language groups mentioned above were included in the final results. Participants received no compensation for their participation, although all of them seemed to enjoy the experience and expressed their excitement about being able to see what happens inside their oral cavity during trombone playing.

7.7.1 NZE participants

Ten trombone players whose native language is NZE were recruited in Christchurch through my personal contacts; unfortunately, I could not finish the analysis of the final participant due to time constraints, and the results from a second, full-length recording session with the pilot participant, collecting both UTI and electromagnetic articulography (EMA) data, could not be included in the final results due to the detrimental effects of the EMA sensors on speech production and tongue position during trombone playing. Table 7.2 below lists basic demographic information for the NZE players included in this study, followed by some relevant individual details.

Participant	Sex	Age	Trombone playing proficiency	Number of years played
S1	m	38	professional	24
S3	m	67	amateur	58
S5	m	33	semi-professional	18
S12	m	61	professional	48
S24	f	25	intermediate	9
S25	m	32	professional	24
S26	m	25	professional	16
S27	m	33	amateur	15
S29	m	68	intermediate	58

Table 7.2: Demographic data for the NZE participants included in this thesis.

The participants in the NZE group included four professional and one semi-professional player, all with a high standard of performance on their instrument; the remaining participants were of amateur and intermediate proficiency (two each). Six players were aged 25-38 while three others were their sixties; overall, participants had played the trombone for an average of thirty years at the time of recording. The lone female player in this group has an English mother and spent two years living in Britain at young age; she also reported elementary proficiency in German and spent more than six months living with a host family in Germany as a teenager. Other interesting information regarding potential non-native influences on speech production concerns S25, who spent seven years living abroad on the West Coast of the USA and in Spain; this participant reported elementary proficiency in both German and Spanish. All other participants in the NZE group were effectively monolingual and never spent significant time outside their home country, although some of them had lived in different locations around NZ.

7.7.2 Tongan participants

I was fortunate to be able to record a single Tongan participant in Christchurch, which provided me with some preliminary findings that promised interesting data if I could make the trip to Tonga to record local players there. A contact with the Royal Tonga Police Band was established through friends of friends on Facebook, and I was able to record all four trombone players employed by the Royal Tonga Police Band, as well as five local amateur players, rounding out my target number of ten participants for each language group. Table 7.3 on the following page lists demographic information for the ten Tongan participants in this study.

Participant	Sex	Age	Trombone playing proficiency	Number of years played
S4	m	43	amateur	28
S14	m	34	semi-professional	19
S15	m	26	semi-professional	10
S16	m	32	semi-professional	17
S17	m	34	professional	17
S18	m	21	amateur	7
S19	m	19	amateur	7
S20	f	26	amateur	8
S21	m	19	amateur	3
S22	m	18	amateur	7

Table 7.3: Demographic data for the Tongan participants included in this thesis.

The make-up of the Tongan group of participants differs markedly from the NZE participants in terms of playing proficiency, the number of years played (12.3 years) and average age (27). Unfortunately, it is not easy to find research participants who play the trombone, speak a specific native language, and are furthermore willing to participate in a study that takes up to two hours; the imbalance of participant groups is a consequence of this situation. A comment is warranted regarding the classification of the Tongan players in terms of their proficiency: although the Royal Tonga Police band is one of two professional music ensembles in Tonga (the other being the Army Band, also a brass band), members also serve as police officers, causing me to rate the playing proficiency of the four band members not strictly based on employment status. The remaining Tongan participants were recruited through contacts provided by members of the Royal Tonga Police Band, and were all graduates of Tongan high schools with music programs. All Tongan participants reported at least elementary proficiency in English as it is a subject in all Tongan high schools. In addition to the participant recorded in Christchurch who had lived in NZ for twenty years, one of the

Police Band players had spent one-and-a-half years living in Brisbane, Australia, while another reported elementary proficiency in Samoan.

7.8 Recording setup

7.8.1 Ultrasound machine

The ultrasound machine used for all recordings was a GE Healthcare Logiq e (version 11), with a 8C-RS wide-band microconvex array 4.0-10 MHz transducer, provided by the NZILBB. Midsagittal video of tongue movements was captured on either a late 2013 15” 2.6 GHz MacBook Pro or a late 2012 HP Elitebook 8570p laptop with a 2.8 GHz i5 processor, both running Windows 7 (64bit); the following USB inputs were encoded using the command line utility FFmpeg (FFmpeg, 2014)[49]: the video signal was transmitted using an epiphan VGA2USB pro frame grabber, and a Sennheiser MKH 416 shotgun microphone connected to a sound devices LLC USB Pre2 microphone amplifier was used for the audio. The encoding formats for video were either the x264 (for video recorded on the MacBook Pro) or mjpeg codecs (for video recorded on the HP Elitebook), while audio was encoded as uncompressed 44.1 kHz mono; frame rates varied between 58 and 60 Hz and were encoded in a progressive scan uyvy422 pixel format at 1024x768.

7.8.2 Using UTI to record trombone players

In order to make it possible to record trombone players who have tubing of their instrument running along the left side of the neck, I suggested modifying a non-metal jaw brace (Derrick, Best, & Fiasson, 2015) previously designed at NZILBB, thereby reducing its width so it would not come in contact with the trombone tubing during playing. The modified probe holder does not extend far enough to contact the trombone tubing except for individuals with a particularly narrow jaw and stabilizes the ultrasound transducer against the jaw, tying probe motion to jaw motion and thus reducing motion variance; a thin strap running along the side of the face does not pose any danger of compromising probe position even if touched by the trombone tubing. As the jaw moves to roughly similar extents during speech production and brass playing (cf. sections 3.1.2 and 3.6.1 above), this introduces uncertainty when

[49] The exact build FFmpeg build used for most recordings and transcoding consisted of the following components: libavutil 52.74.100; libavcodec 55.57.100; libavformat 55.36.101; libavdevice 55.11.100; libavfilter 4.3.100; libswscale 2.6.100; libswresample 0.18.100; libpostproc 52.3.100]

comparing tongue position measurements across time. However, an assessment of the motion variance of the system, evaluating tongue and head movement data collected using both UTI and EMA (Derrick, Best, & Fiasson, 2015), showed that 95 percent confidence intervals of probe motion and rotation were well within acceptable parameters described in a widely-cited paper that traced head and transceiver motion using an optical system (Whalen et al., 2005). Alternative systems exist that either completely fix the ultrasound probe in place (Stone & Davis, 1995; Articulate Instruments Ltd., 2008) or measure, and thus allow to correct for, head motion (Mielke et al., 2005; Miller & Finch, 2011); however, either system would have been unsuitable to record trombonists due to the positioning of the instrument next to the head. Furthermore, while a rigid system such as the helmet by Articulate Instruments Ltd. (2008) may have less short-term motion variability than the jaw brace used for this research, Derrick, Best, and Fiasson's (2015) data appears to show that this system outperforms the head stabilization helmet in terms of "long-term slippage" (Derrick, Best, & Fiasson, 2015, p. 4).

7.8.3 Microphone placement for audio recordings

Ultrasound machines usually include their own CPU, which means that they often emit considerable fan noise that can present problems for researchers recording speech in an enclosed space. Some researchers have addressed this issue by placing the ultrasound machine in a separate room and feeding the cable connecting the ultrasound transducer through an opening in a shared wall; such an approach, however, was not possible for this research. Instead, we placed the microphone as close as possible to the participants' lips for the speech elicitation task, significantly reducing the amount of fan noise on the recordings. For trombone recordings such close microphone placement is not advisable due to its effect on the recorded sound spectrum; Benade et al. (1977) showed that the spectrum recorded by a microphone placed "one bell radius in front of a brass instrument bell" closely approximates the result "obtained by spectral averaging of sounds recorded at many points in the room" (p. S35). This approach was followed for the recordings described in this thesis, however, distance was not strictly controlled for. Rather, participants were asked to keep the same distance to the microphone throughout the music part of the recording session, although they sometimes moved closer or further away while playing. Unfortunately, it would have been impossible to directly attach the microphone to the trombone bell without adding considerable weight to the instrument; furthermore, I considered it important to avoid restricting the participants in their natural movements during playing as they were already to a highly unusual playing situation with the ultrasound probe held in place underneath their jaw.

A mishap in determining the FFmpeg command line code for recording ultrasound videos on the HP laptop used for the trip to Tonga meant that the laptop’s internal microphone was selected as audio input instead of the shotgun microphone connected to the USBPre2, resulting in insufficient audio quality as the laptop was positioned quite close to the ultrasound machine’s fan; this also affected the first participant recorded in Christchurch after the trip. Luckily, the camcorder used to record video of the players’ face recorded reasonable audio quality, and this audio stream could be used to carry out the analysis for these participants; synchronization was carried out using the /tatatatata/ (or /dadadadada/) syllables elicitated at the beginning and end of each of recording block, as detailed in section 8.2. Naturally, the quality of this audio stream does not quite approach that of data recorded using the correct settings, further limiting the possibility of audio data comparisons across participants, already compromised by recording in different rooms.

7.9 Recording procedure

A typical recording session for this thesis would proceed in the following fashion: participants would have been given a short summary of the project (avoiding disclosure of any details regarding findings and hypotheses) when approached about participating in the study. They were told that the recording session would last up to two hours, although the whole process usually did not take longer than one-and-a-half hours. When arriving at the recording location, participants were first asked to read through the information sheet and sign the corresponding consent form, as well as completing a short questionnaire to collect basic demographic data and some information about their playing experience and potential Second Language proficiency (copies of the Human Ethics application and corresponding documents can be found in appendix B, pages 255-263). After filling in the relevant documents, participants were given a few minutes’ time to warm up on, and familiarize themselves with, the ‘pBone’ and the provided mouthpiece. The mouthpiece had previously been cleaned with isopropyl alcohol while participants were in attendance (in accordance with Human Ethics regulations), and the whole instrument was cleaned with warm water before every recording session to minimize the effect of particles accumulating within the bore.

Subsequently, the experiment proper would begin with participants being asked to put on the jaw brace with the ultrasound transducer, and adjustments were made to allow for a comfortable fit. In most cases, ultrasound gel was also applied to improve the ultrasound image quality, in addition to adjusting the settings available on the ultrasound machine. Around the same time, participants would be informed about the typical sequence of events for each block of wordlist elicitations, especially the /tatatatata/ or /dadadadada/ syllables to be uttered at the beginning and end of each block to allow correcting for audio and video misalignment. Usually, I would also ask participants to slowly swallow water to record a first palate trace (cf. Epstein & Stone, 2005), which, at the same time, served the purpose of checking whether the ultrasound scan depth was set correctly. Additional palate traces would be recorded throughout the recording procedure, especially after starting the trombone playing part of the procedure which carried an additional risk of altering the position of the jaw brace. A video recording of the participant's upper body and face was made during the actual data collection using a full-HD Sony Camcorder (model number HXR-MC50E) to allow tracking any possible changes in jaw brace/probe position, and to document when palate traces were recorded; this recording was started at this time.

The first part of the experiment consisted of six to nine blocks of wordlist readings, depending on the native language of the participant. Tongan participants were asked to read three blocks of a shortened list of English words in addition to the Tongan words; however, this data was not analyzed after it became clear that none of the participants produced native-like pronunciations, diminishing the possible impact of their English language proficiency on the experimental results. Following the completion of this first part of the recordings, participants were offered to take a short break while an myself or another person assisting with recordings moved a switch on the USB Pre2 audio device implementing a -15 dB reduction, and replaced the laptop presenting the wordlist stimuli with the sheet music for the musical passages. All participants stood for the duration of the actual recordings as this represents the most comfortable position of playing the trombone, given the circumstances.

Before starting the second part of the experiment, participants were reminded to avoid touching the jaw brace with the instrument tubing, and given a short time to play a few notes under these conditions (if I had not given them time to do so earlier). Prior to starting the recording of each musical passage, a pre-specified tempo was indicated to the participants using a metronome (the metronome was not kept on during the recording); any special instructions were also given at this time and participants were reminded to utter the syllables /tatatatata/ or /dadadadada/ at the beginning and end of each recording. After completing the recordings, participants were given the chance to face the screen of the ultrasound machine and observe their own tongue movements during playing. Quite often participants were puzzled by what they saw, indicating that they had no idea that this was what they were doing with their tongue.

8 Data analysis

8.1 Video transcoding

The first step of the analysis involved transcoding the videos recorded with the x264 codec to the mjpeg codec, using FFmpeg (FFmpeg 2014). This step was necessary to account for hardware limitations of the computer used to trace the ultrasound images in MATLAB. The x264 codec stores image data in a highly compressed format, which reduces file size but requires a fast CPU to play back, while the mjpeg codec stores every video frame as a single jpeg image, leading to a large file size but reducing hardware requirements during playback or when accessing single frames. Subsequently, a copy of the audio track was spliced off from the video so it could be used to carry out audio-video alignment, and to perform automatic speech segmentation using the LaBB-CAT tool (see section 8.3.1 below).

8.2 Audio-video alignment correction

Since I recorded two different USB inputs at the same time, the audio track always preceded the video track by thirty to a hundred milliseconds. This misalignment probably resulted from the much larger data throughput of the frame grabber, which placed high demands on the recording laptop either in terms of encoding/compression (x264 codec) or the amount of data to be kept in the RAM and written to the hard drive (mjpeg codec). In order to enable correcting for this difference, participants were asked to produce /tatatatata/ (or /dadadadada/) syllables at the beginning and end of each of recording block (cf. sections 7.4.1 and 7.9 above). This allowed me to identify the exact video frame when the tongue first started dropping down from its place of articulation with a sharp rise in the audio waveform (see Miller & Finch, 2011), using the annotation software ELAN (Wittenburg et al., 2006); more specifically, I measured the time difference between the two events described above and added a corresponding duration of silence to the beginning of the relevant audio track in Praat (Boersma & Weenink, 2014) before verifying the result. Unfortunately, the time lag between the two tracks differed even for consecutive recording blocks, rendering this a time-consuming, but absolutely necessary, activity.

8.3 Speech segmentation

8.3.1 Speech segmentation using the HTK toolkit implemented in LaBB-CAT

After fixing temporal alignment, a transcriber transcript file (Barras et al., 2001) matching the audio track was created using a Java application by LaBB-CAT developer Robert Fromont, which required the input of a text file containing the stimuli and their distribution into blocks. This process was simplified by using the same sequence of words for all NZE and Tongan participants, respectively. The transcript file was then added to a corpus of participant data created in LaBB-CAT (Fromont & Hay, 2012), and the speech audio was automatically segmented using the HTK tool (Young et al., 2006). The phonemes matching the orthography of the input were exported from the AE version of the CELEX2 dictionary (Baayen, Piepenbrock, & Gulikers, 1995) for the NZE stimuli, while a custom dictionary was created for all the words contained in the Tongan wordlist. As the segmentation process for the Tongan data relied on an algorithm developed for speech in English, the results were expectedly worse than for the NZE data, requiring extensive manual correction. A final step of the segmentation process automated in LaBB-CAT involved creating an additional layer of phoneme encoding in X-SAMPA (Wells, 1995), before files were exported as Praat transcripts (Boersma & Weenink, 2014).

Due to systematic differences of the phonologies of AE and NZE, a few NZE phonemes required manual corrections, and I manually checked all exported textgrids using common acoustic landmarks to make sure segmentation was performed correctly. For vowel duration, I went with the presence of a clear F2 and its relative steadiness; some of the NZE monophthongs (FLEECE - cf. section 7.3.1, THOUGHT) were prone to be realized as diphthongs, particularly by younger speakers, so that I only selecting the steady portion of these vowels (excluding on- and/or off-glides). I also added an additional tier which kept the full phonemic transcription (in X-SAMPA) of the word but marked the selected vowel portion by inserting a dash (-) before and an equals sign (=) after the respective phoneme to enable searches using regular expression later on in the analysis process; an example for what such a label would look like is *g-'6:=d@n* ('garden'). For both the NZE and Tongan data, only primarily stressed (or accented) vowels were selected for analysis; for the NZE data I used the stress markings from the New Zealand Oxford Dictionary (Kennedy & Deverson, 2005) entries, while I applied the penultimate stress/accent rule to the Tongan data, as outlined in section 7.3.2.

8.3.2 Manual segmentation for early data

When first starting the analysis of participant data, I manually segmented the speech data for the first two NZE participants (S1 – pilot data – and S5) and the first Tongan participant (S4). For the NZE data, I used phonemic transcriptions from the New Zealand Oxford Dictionary (Kennedy & Deverson, 2005), although I substituted the more phonetically accurate IPA symbols suggested by Bauer and Warren (2004), and added schwa (/ə/) to distinguish this phoneme from the stressed KIT vowel (/ə/). These transcriptions were also used to manually check the phonemic transcriptions in the X-SAMPA layers exported from LaBB-CAT (see section 8.3.1 above). The same acoustic cues as outlined below for manually checking the automatically segmented data were used to segment the speech signal. For the Tongan data, I had to rely on the phonemic correspondence of pronunciation to spelling, and the phonemic transcription of the manually annotated dataset was used to create the 'dictionary' used by LaBB-CAT to assign phonemes to the speech produced by the other Tongan participants.

8.4 Annotation of musical passages

For the musical passages, I used the Praat 'Annotate - to TextGrid (silences)' tool to perform a rough segmentation of the audio signal into different notes, manually corrected the boundaries, and then used a script to assign the appropriate label to each note (Boersma & Weenink, 2014). Each label included the following information (if relevant): number of the musical passage, measure, dynamic marking, pitch (followed by *.n* if a measure contained more than instance of the same pitch), and note duration if shorter than a crotchet/quarter note; here is an example: *4e.7 mf acc Bb3.4*. Missed notes were eliminated, although for long sustained notes, I used a later part of the note if the participant recovered to produce a well-formed note; a few of the less proficient participants sometimes produced incorrect pitches, which I relabeled to reflect the played note given that a participant otherwise executed the musical passage correctly.

8.5 Selection of the temporal location for ultrasound images to be analyzed

After double-checking the speech and music annotations, I ran a Praat script (Boersma & Weenink, 2014) on each textgrid that created a point tier with points at the midpoint of vowels, and at one-third of interval duration for the trombone notes; the script also copied the labels from the corresponding interval tier. The only exception to this procedure was that I manually selected an appropriate steady state of tongue position during note production for the pilot participant, before realizing that performing this process for all participants would be too time-consuming. The decision to use one-third of note duration for the other participants was based on the observation that participants often reduced the intensity of notes after the initial attack, resulting in a longer ‘tail’; using one-third of note duration also guaranteed that the tongue position would be relatively stable following the tongue movement associated with the articulation of notes.

8.6 Tongue contour tracing using GetContours in MATLAB

The GetContours script (Tiede & Whalen, 2015) by Mark Tiede for MATLAB (Mathworks Inc., 2015) was used to trace the midsagittal ultrasound tongue contour in each of the video frames identified using the steps outlined above. This script exports timestamps and annotations from a specific tier in a Praat textgrid (in this case a point tier), and pulls up the respective frame in a figure window; one then places at least three so-called anchor points outlining the produced tongue shape by clicking on top of the image. These anchors are automatically connected by a cubic spline fit, making it easy to visually confirm whether they accurately represented the tongue contour in the underlying ultrasound image; the script also writes all relevant information (time stamp, frame number, and copy of the image, label, anchors, curve shape interpolated to one hundred points) to a variable in the MATLAB workspace. In a few cases, MATLAB had problems dealing with corrupted video frames, preventing me from tracing the relevant frame, and more seriously, advancing past this frame to trace later ‘key frames’ in the same video; fortunately, Mark Tiede was able to provide a quick fix for this issue, meaning that only a few tokens for Tongan participant S4 were lost.

8.7 Outlier removal, data transformations and export from MATLAB

8.7.1 Outlier removal

Once all vowel or note tokens were traced for a certain participant, I employed regular expressions to assemble all tokens for a specific stressed/accented vowel or note into a variable, which I then used to plot all tokens and remove visual outliers. At this stage, I also used the palate traces, estimated by tracing the trajectory of the tongue along the roof the mouth in a sequence of images recorded during swallowing (cf. Epstein & Stone, 2005), to check for possible movements of the probe holder. Although the reliability of the palate traces themselves was quite low due to variable image quality and the potential effect of swallowing movements on ultrasound transducer position (via jaw opening), comparing the palate traces for NZE participant S12 helped confirm that the probe holder had moved during recordings. Further evidence arose from examining video of the participant's face showing that he bumped the ultrasound holder with the tubing of the 'pBone'. To correct for this unwanted change in transducer position, I rotated and shifted the traces recorded after the mishap in relation to the estimated transducer position (see following section) until the palate traces and the overall area outlined by the traces roughly matched up; naturally this means that the overall reliability of S12's sustained note data is somewhat compromised. Figure 8.1 on the following page shows S12's tongue and palate traces prior to, and after, completing this process. Unfortunately, a similar mishap occurred with the pilot participant's data, meaning that I had to discard a substantial number of vowel tokens as I had not recorded any palate traces that would have enabled me to correct for unwanted transducer movement.

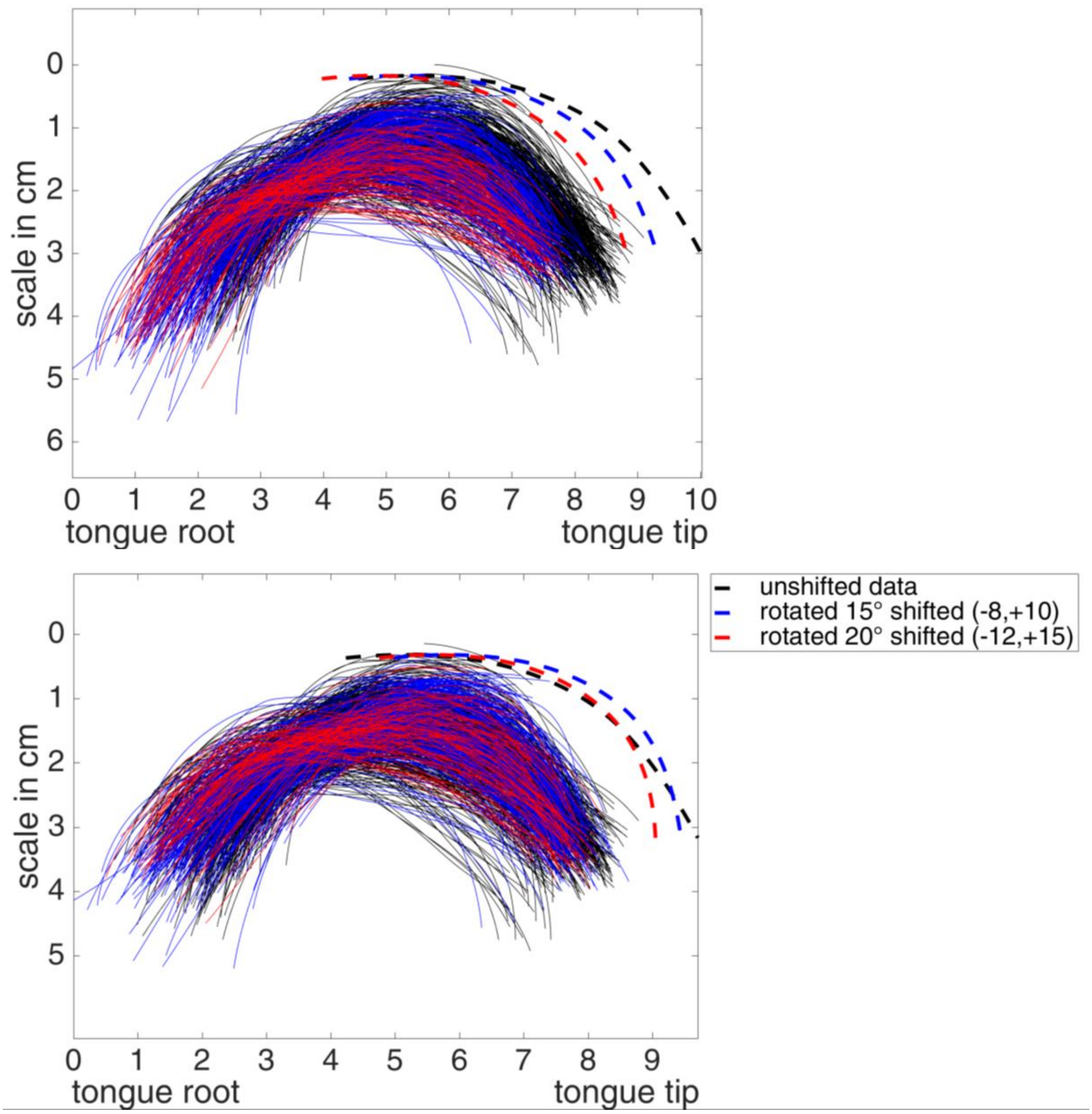


Figure 8.1: Tongue and palate traces for participant S12 NZE prior to (top plot) and after correcting for unwanted ultrasound transducer movement (bottom plot).

8.7.2 Estimation of 'virtual origin' and transformation of data to polar coordinates

The estimated transducer origin, which I will refer to as 'virtual origin' hereafter, plays a pivotal role for all subsequent steps of the analysis, as it provides a principled way of comparing UTI data across participants due to its physical relationship with the recorded images. For further motivations of using the virtual origin to transform radial UTI data to polar coordinates, and the necessity of using polar coordinates to calculate

SSANOVA average curves of ultrasound traces, see Heyne and Derrick (2015c) and section 8.8 below.

Figure 8.2 below shows how the transducer position can be estimated from any radial ultrasound image by overlaying two lines, and setting their equations equal to determine the origin coordinates; the figure also shows several horizontal lines that were overlaid to estimate the conversion rate to convert the data to mm for each participant (GetContours exports coordinates based on image pixels).

The virtual origin was subsequently used to transform the one hundred interpolated points of each ultrasound trace to polar coordinates, employing the function cart2pol in MATLAB (Mathworks Inc., 2015); as part of this step, all coordinates were normalized to the virtual origin defined as (0, 0), and a separate dataset was created with the data transformed to a mm scale (in Cartesian coordinates) based on the conversion rate estimated as detailed above.

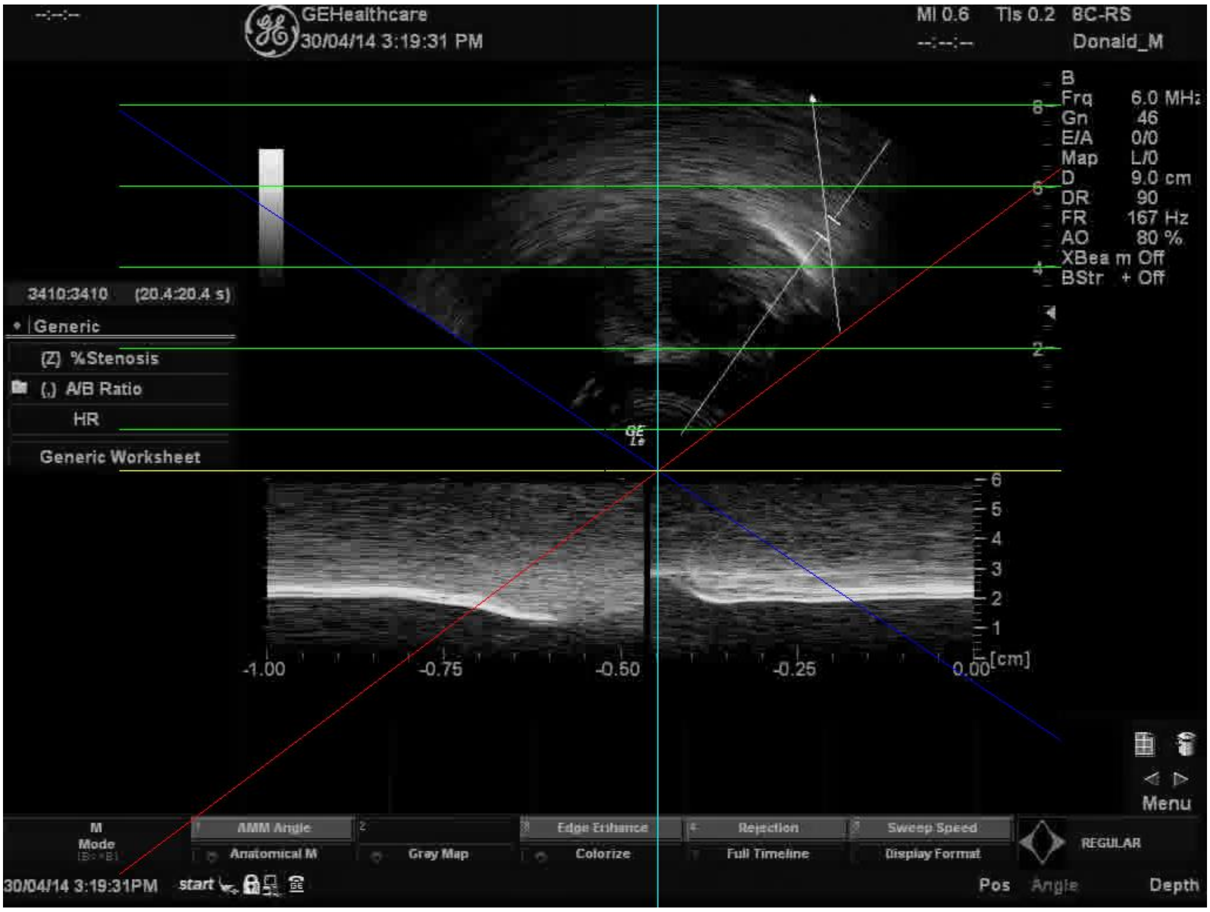


Figure 8.2: Estimation of the ‘virtual origin’ and pixel scale from a randomly selected ultrasound image by overlaying various lines.

8.7.3 Cutting traces to avoid edge effects on average curves

Due to variable image quality and the unconstrained placement of GetContours anchors on each ultrasound frame, individual tokens differed greatly in length (cf. figure 8.1 above). To avoid that a single trace would determine the shape of an average curve calculated for all tokens of a certain sound or sustained note at the edges, I wrote a MATLAB script that eliminated points with extreme angle values (the Theta value for polar coordinate pairings) when less than twenty percent of combined traces for vowel or sustained note tokens provided any points below or above these values. This has the effect of cutting off traces past imaginary fan lines extending from the virtual origin, as illustrated in figure 8.3 below.

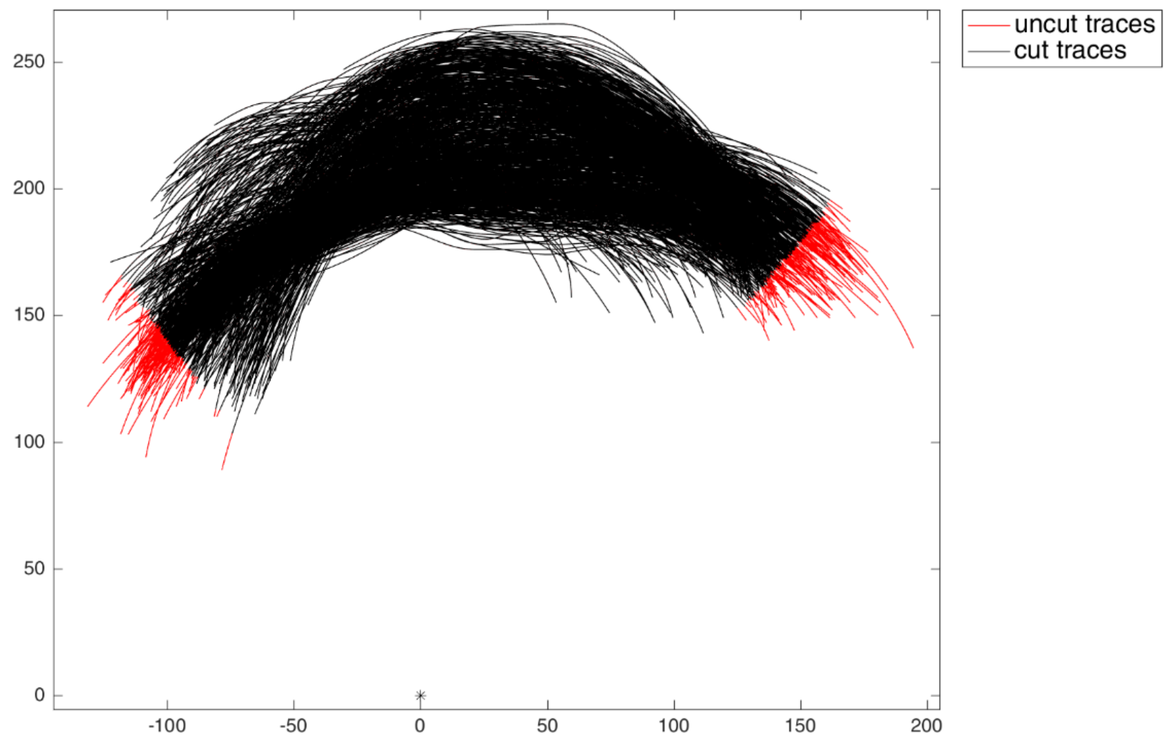


Figure 8.3: Image illustrating the cutting off of extreme values along fan lines extending from the virtual origin for S29 NZE; the scale is in pixels and normalized to the virtual origin (0, 0).

After extreme values were eliminated according to this process, all traces were re-interpolated to one hundred points, and exported from MATLAB as comma-separated values (.csv). Tables 8.1a&b and 8.2a&b below list the final token numbers for all participants, with these observations forming the basis for the average calculations detailed in the following section. Overall, I manually traced 18,521 ultrasound images, in addition to the participant data that did not make it into this thesis, and the consonant and place of articulation data presented in section 9.5.4 below.

	/ɐː/	**/ɐ/**	**/e/**	**/ɛ/**	**/iː/**	**/ɘ/**	**/oː/**	**/ɒ/**	**/ʉː/**	**/ʊ/**	**/ɵː/**	**/ə/**	**/ə#/**	*SUM*
S1	7	8	8	5	6	7	8	3	5	3	4	7	1	72
S3	46	61	57	61	52	72	34	50	49	17	58	218	87	862
S5	57	70	69	61	56	87	32	64	56	21	58	171	80	882
S12	55	64	56	55	49	74	35	51	50	19	48	114	72	742
S24	61	68	59	74	60	82	39	59	53	22	59	201	91	928
S25	61	69	66	77	59	83	37	59	49	21	64	154	84	883
S26	59	69	43	65	37	75	23	61	34	21	55	156	89	787
S27	60	73	50	76	35	81	32	57	44	25	62	149	92	836
SUM	464	550	470	554	415	643	276	466	389	172	471	1374	686	*6930*

Table 8.1a: Final numbers of vowel tokens for all NZE participants.

	Bb2	**F3**	**Bb3**	**D4**	**F4**	**low notes**	**middle notes**	**high notes**	*SUM*
S1	26	47	44	23	13	73	44	36	153
S3	84	164	160	63	17	248	160	80	488
S5	80	157	139	55	18	237	139	73	449
S12	74	145	131	47	16	219	131	63	413
S24	32	95	71	30	12	127	71	42	240
S25	76	150	134	49	16	226	134	65	425
S26	79	149	124	49	16	228	124	65	417
S27	82	146	130	37	11	228	130	48	406
SUM	609	1200	1065	406	135	1809	1065	541	*3415*

Table 8.1b: Final numbers of sustained note tokens for all NZE participants.

	/a/	/e/	/i/	/o/	/u/	*SUM*
S4	79	35	54	37	24	229
S14	132	78	106	94	68	478
S15	123	73	103	90	69	458
S16	114	76	105	83	54	432
S17	131	79	112	95	71	488
S18	116	76	108	90	63	453
S19	127	71	113	92	69	472
S20	129	70	100	97	60	456
S21	129	73	112	88	68	470
S22	133	79	112	92	71	487
SUM	1213	710	1025	858	617	*4423*

Table 8.2a: Final numbers of vowel tokens for all Tongan participants.

	Bb2	**F3**	**Bb3**	**D4**	**F4**	**low notes**	**middle notes**	**high notes**	*SUM*
S4	44	96	85	25	8	140	85	33	258
S14	63	142	134	50	17	205	134	67	406
S15	74	147	127	43	16	221	127	59	407
S16	82	160	143	44	16	242	143	60	445
S17	72	148	128	47	18	220	128	65	413
S18	69	129	112	48	16	198	112	64	374
S19	37	112	105	43	16	149	105	59	313
S20	57	149	142	58	13	206	142	71	419
S21	63	118	106	46	12	181	106	58	345
SUM	626	1345	1190	448	144	1971	1190	592	*3753*

Table 8.2b: Final numbers of sustained note tokens for all Tongan participants.

8.8 Calculation of SSANOVA average curves in R

The free software environment for statistical computing and graphics, R (R Development Core Team, 2008), was used to calculate average tongue contours by token (different vowels and sustained notes) and participant, based on the traces imported from MATLAB (Mathworks Inc., 2015). The statistical technique used to calculate these average curves is called smoothing spline analysis of variance (SSANOVA; cf. Gu, 2013b), which fits a natural cubic smoothing spline ("a piecewise polynomial function that connects discrete data points called knots") based on the variation in the provided data (Davidson, 2006, pp. 409-410). Fitting this curve includes two terms which are numerically minimized,

> one that attempts to fit the data and one that penalizes a fit which does not have the appropriate amount of smoothness. Although the penalty term does not allow the function to fit the data precisely, it ensures that the resulting spline has a suitable amount of smoothness (Davidson, 2006, p. 410).

In addition to the average curve fit, usually estimated for a number of equally spaced x values, the SSANOVA technique also provides an estimate of the standard error at each point location, which can be plotted as error bounds around the average curve by adding (upper bound) and subtracting (lower bound) the standard error multiplied by 1.96 to the estimated value. The package used to perform these calculations in R is called gss (General Smoothing Splines; Gu, 2013a).

After I had calculated SSANOVA average curves for the pilot participant, I realized that basing these calculations on data points in the Cartesian plane "creates errors that are compounded at tongue tip and root where average tongue shape deviates most from a horizontal line" (Heyne & Derrick, 2015c, p. EL509). These errors were especially pronounced for a comparison of averages based on a small number of vowel tokens and a large number of sustained note tokens, almost nullifying the patterns shared by both kinds of tokens, and easily recognizable when plotting the traces together. Converting the data to polar coordinates solved this issue, and although I later found an R script by Mielke (2013) that performs this transformation to polar space using a polar coordinate system origin based on extreme values within the dataset (cf. Mielke, 2015), using the estimated transducer position as origin provides a more principled way of transposing the data, which we published in the *Journal of the Acoustical Society of America - Express Letters* (Heyne & Derrick, 2015c).

Figure 8.4 on the next page shows the underlying traces and SSANOVA averages curves, as well as their error bounds (barely visible), for all tokens of selected vowels and sustained notes produced by participant S29 (NZE). I used the ggplot2 package (Wickham, Chang, & RStudio, 2016) for all plots in this thesis that were produced using the R environment; the previous figures in this section, however, were created in MATLAB.

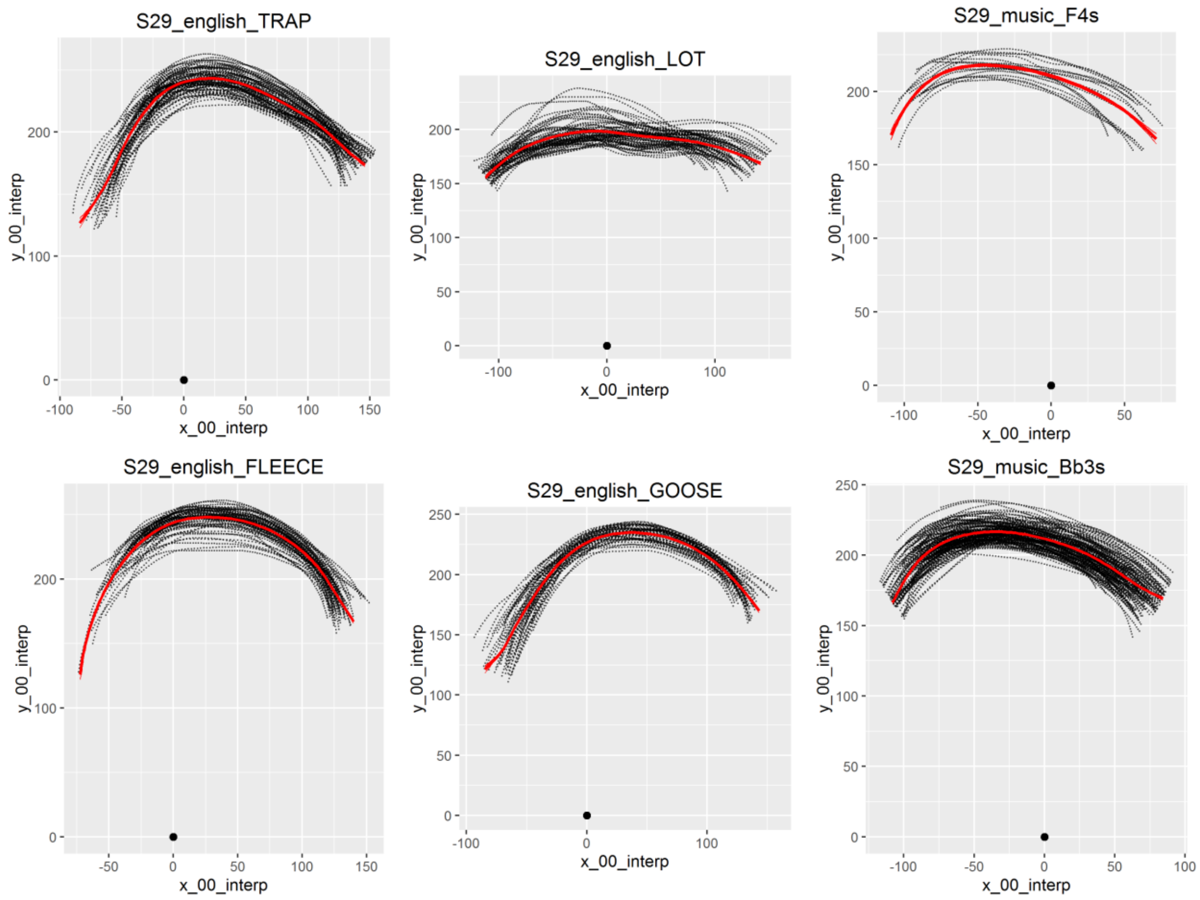


Figure 8.4: SSANOVA average curves with error bounds (red) and underlying traces (black) for selected vowels and notes produced by S29 NZE; the virtual origin is shown by a point at (0, 0).

8.9 Z-scoring of data

To answer the question whether a player's native language influences their sustained note production on the trombone, I had to carry out direct comparisons of the articulatory data collected for both language groups investigated in this study, and thus, across individuals. Very few researchers using UTI have attempted such inter-individual comparisons, which are complicated by the lack of anatomical landmarks imaged by the technique, in addition to individual differences in vocal tract shape and biomechanics (cf. section 3.3.4 above). Furthermore, the insufficient quality of the collected palate traces and the omission of collecting other possibly helpful data, such as an estimate of the occlusal plane using a bite plate (Scobbie et al., 2011), means that no anatomical landmarks were available for the data collected for this thesis.

A few researchers, however, have suggested methods for determining and comparing, e.g., the curvature of selected tongue shapes (Ménard et al., 2011; Stolar & Gick, 2013; Zharkova, 2013a&b; Dawson, Tiede, & Whalen, 2015) or the relative articulatory height and fronting of a certain vowel tongue shape (Lawson & Mills, 2014; Lawson, Mills, & Stuart-Smith, 2015), independent of anatomical landmarks. Still, none of these methods allow for the unbiased comparison of UTI measurements from two different activities such as speaking and trombone playing, for which it is unclear whether overall tongue shape or the distance between articulatory targets is more important.

The solution to address the problem in this thesis arose out of the same considerations that motivated the selection of the virtual origin as origin of a polar coordinate system for calculating SSANOVA average curves of midsagittal ultrasound data. The central assumption is that the location of all the points exported by tracing the tongue contour on individual ultrasound frames is relative to the virtual origin, which makes it possible "to rotate individual traces and/or average curves without affecting the true variation underlying their error estimates" (Heyne & Derrick, 2015c, p. EL513). It was thus necessary to identify an articulatory gesture produced by all participants and constrained as much as possibly by each individual's anatomical and biomechanical properties rather than language-dependent factors; these criteria are satisfied by the high front vowel /i(ː)/ (FLEECE in NZE), which can be presumed to follow similar articulatory constraints across both language groups. The highest point of the tongue and its horizontal offset in relation to the virtual origin thus provide a measurement that can be z-scored across participants, allowing inter-individual comparisons. Note, however, that such an approach necessarily ignores individual differences concerning palate shape (doming) and length, and other features further down the vocal tract; furthermore, differences in the size of male and female articulatory spaces could distort the results, even though the make-up of the dataset should account for this potential confound, with only one female participant in each language group (cf. section 3.3.4 above). Finally, the placement of the ultrasound transducer in relation to the articulatory space mapped out by a participant's tongue movements can differ across participants (cf. Heyne & Derrick, 2015c, p. EL 513) and cannot be accounted for by the z-scoring procedure as its main advantage lies precisely in maintaining this relationship.

Following these considerations, I calculated the highest point of the tongue for the SSANOVA average curves representing each participant's high front vowel in relation to the virtual origin (in the case of the pilot participant the highest vowel tongue position was DRESS (/e/), conceivable for an 'advanced' speaker of NZE; cf. Maclagan & Hay, 2004, p. 187); using the points transformed to polar coordinates, this value is provided by the maximum radial length (Rho) and its horizontal position in relation to the virtual origin is given by the angle value (Theta). To preserve the underlying precision for at least one dataset, the polar coordinates representing the highest point of the tongue for the participant with the smallest vocal tract (corresponding to the smallest scan depth setting) was chosen to provide the basis for z-scoring all other participants' data. Figure 8.5 on the following page illustrates how this process led to the rotation and/or scaling of various participants' data; note that for all participants except the pilot data (S1), a display of anatomical m-mode data was situated beneath the wedge-shaped area showing the full midsagittal tongue shape, leading to a big difference in the required scaling (cf. figure 8.2 above). A table listing the required amount of rotation and z-scoring for all participants can be found in Appendix C, p. 265.

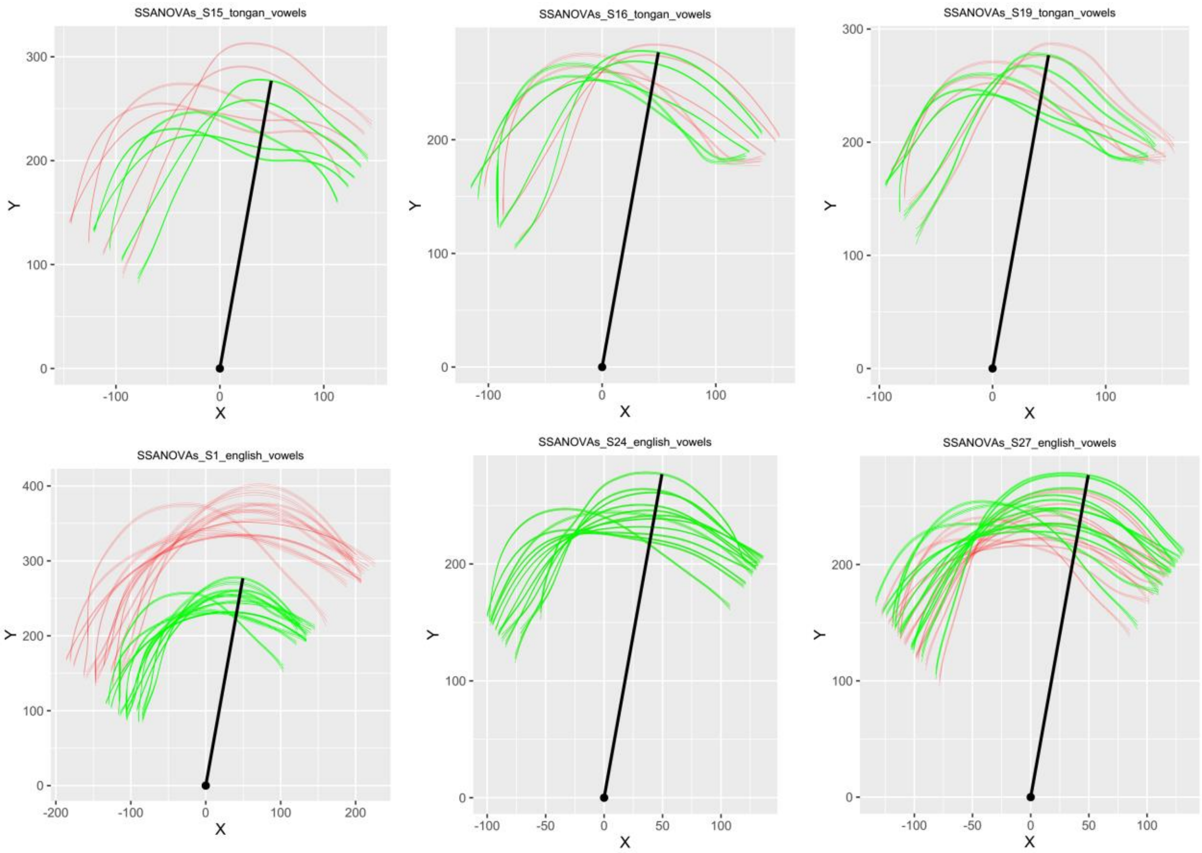


Figure 8.5: Plots illustrating the effect of the z-scoring procedure on the SSANOVA average curves for monophthong productions by three Tongan (top row) and three NZE participants (bottom row). S24 (middle of bottom row) featured the smallest scan depth setting and was thus chosen to provide the target vector for the z-scoring procedure, shown as a black radial line extending from the virtual origin in each plot.

8.10 Chapter summary

In this chapter, I have described the various steps involved in analyzing the midsagittal ultrasound video collected for this thesis. After several smaller steps involving data pre-processing, speech segmentation, and annotation of the musical passages, I explained how I selected the ultrasound images that I subsequently traced in MATLAB using the GetContours script. Analysis proper only began at this stage; while outlier removal was relatively trivial yet time-consuming, discovering a principled way for averaging and normalizing the large amount of ultrasound data collected for this thesis was a tedious process. I am nonetheless confident that I have found a reliable way of doing so that can be applied in future studies employing midsagittal UTI.

9 Results

This section presents the results of the measurements carried out according to the analysis procedures outlined in the previous chapter. Since UTI does not provide any landmark features of the vocal tract, the findings are determined by comparing the measurements for various tongue positions to one another, which in this case consist of smoothing spline ANOVAs fit on large numbers of tokens.

9.1 Speech data

Figures 9.1 and 9.2 show the pooled midsagittal tongue contours for the two language groups investigated in this thesis, averaged across all participants (n=9 for NZE and n=10 for Tongan); as is conventional for articulatory research, the front of tongue is shown at the right while the back of the tongue is at the left. Note that the SSANOVA average curves are behaving erratically outside the area enclosed by the overlaid fan lines extending from the estimated ultrasound transducer position, due to the varying length of the underlying z-scored tokens (in turn caused by variable image quality at the edges, cf. chapter 6 above); these extreme areas were thus excluded from the numerical comparisons reported below. The identity of the various monophthongs of NZE is indicated in the legend by the use of IPA symbols as outlined in sections 7.3.1 and 8.3.2 above (for a mapping to the lexical sets by Wells (1982) see figure 7.2); for Tongan, the standard IPA symbols for five-vowel systems are used. The scale for both plots is identical, although it is arbitrary due to being based on the participant with the smallest vocal tract; similar colors are used across the two languages to indicate vowels that are acoustically similar (or bear a relationship based on their historical pronunciation as in the case of the NZE GOOSE vowel).

Figure 9.1 on the next page shows the average curves for the five monophthongs of Tongan, based on the productions of ten speakers. Although Tongan distinguishes short and long vowels (often analyzed as one or two morae, respectively, cf. section 7.3.2 above), the articulatory differences between these phonemes are very small, and I thus decided to treat these tokens as a single underlying motor target.

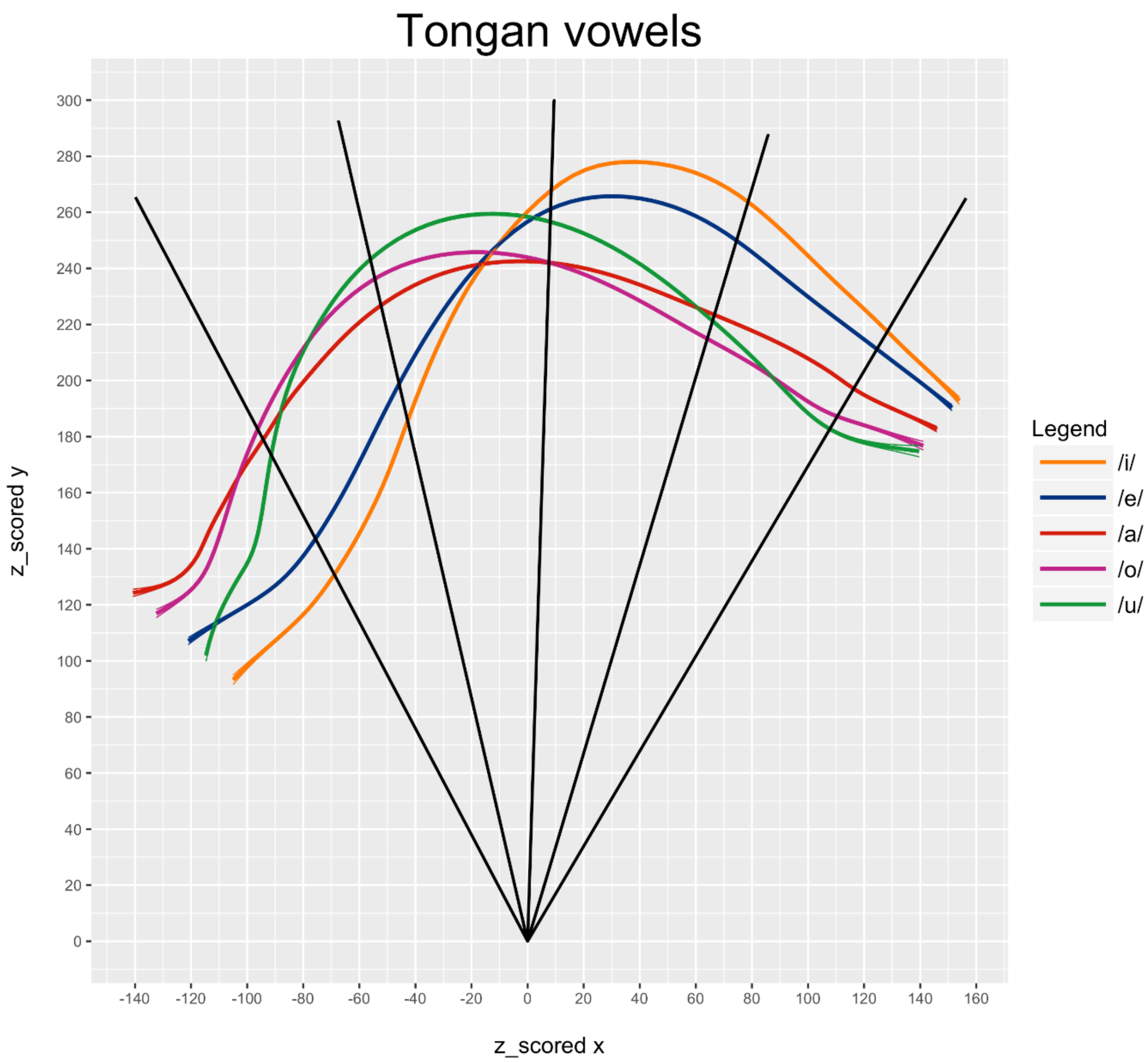


Figure 9.1: SSANOVA average curves for the z-scored tokens of the five vowels of Tongan (including short and long phonemes), produced in accented position and averaged across the articulations of ten participants. The front of the tongue is to the right of the image while the back of the tongue is shown at the left. 95 percent confidence intervals are plotted as upper and lower bounds around the SSANOVA average curves (using the same colors), even though they are barely visible aside from at the edges. The overlaid fan lines (in black) indicate the estimated ultrasound transducer position.

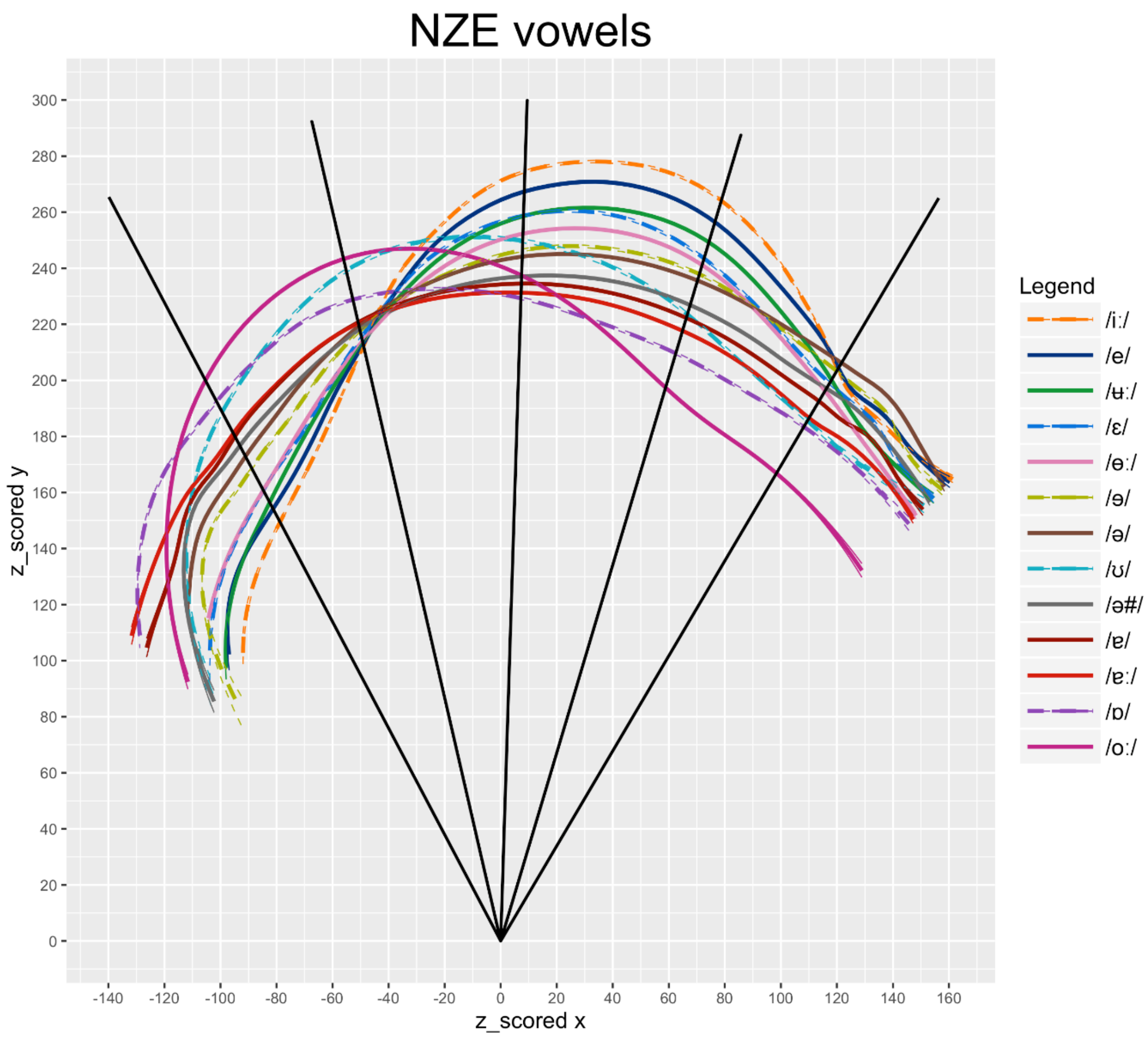


Figure 9.2: SSANOVA average curves for the z-scored tokens of the stressed monophthongs of NZE plus schwa in non-final and final position, averaged across the productions of nine participants. The front of the tongue is to the right of the image while the back of the tongue is shown at the left; the legend indicates vowel quality using the IPA symbols suggested by Bauer and Warren (2004; plus schwa /ə/). 95 percent confidence intervals are plotted as upper and lower bounds around the SSANOVA average curves (using the same colors), although they are barely visible aside from at the edges. The overlaid fan lines (in black) indicate the estimated ultrasound transducer position.

Figure 9.2 above on NZE matches up well with formant plots for moderately ‘advanced’ speakers of NZE reported in the literature (cf. section 7.3.1 above), although numerous studies (e.g., Scobbie, Stuart-Smith, & Lawson, 2012) have shown that the mapping of formant values onto articulatory gestures/positioning is of course far from straightforward, particularly for back and low vowels (cf. Esling, 2005). The DRESS vowel (/e/) is characteristically close, accompanied by a close TRAP vowel (/ɛ/) and a centralized KIT vowel (/ɘ/); furthermore, the GOOSE vowel (/ʉː/) is quite fronted (see Lawson, Mills, & Stuart-Smith (2015) for an articulatory and acoustic study of GOOSE fronting in different varieties of British English). Note that the DRESS vowel (/e/) is articulatorily higher than the centralized KIT vowel (/ɘ/), although this is not the case acoustically (see Heyne & Derrick, 2016a). In addition to monophthongs occurring in stressed position, the figure also includes two different average curves for unstressed schwa occurring in non-final and final position. This sub-phonemic difference has been reported for a number of varieties of English (see various chapters in Foulkes & Docherty, 2014; Mesthrie, 2008 on African Englishes; Penhallurick, 2008 on Welsh English), but I am not aware of any study that has investigated the phenomenon systematically (but see Yamane-Tanaka, Gick and Bird (2004) for a small study of functional versus lexical schwa). For NZE, we have recently been able to show that this difference holds for a large number of speakers reading from a wordlist in the most recent corpus from the Origins of New Zealand English project (ONZE; Gordon, Maclagan, & Hay, 2007), and interacts with speaker year of birth (Heyne & Derrick, 2016c); however, it does not seem to be socially marked as in ‘New Australian English’ (Kiesling, 2005). Even though this thesis is generally concerned with how phonemic differences affect the motor movements used by speakers of a certain language when playing brass instruments (and thus the *phonology* of the two chosen languages), the sub-phonemic distinction of schwa occurring in different environments in NZE was included based on its substantial articulatory distance, and the high frequency of occurrence of reduced vowels in speech.

9.2 Music data

Figures 9.3 and 9.4 on the following pages show the average tongue contours for the sustained productions of five different notes played on the trombone, superimposed on the average vowel productions for the respective languages.

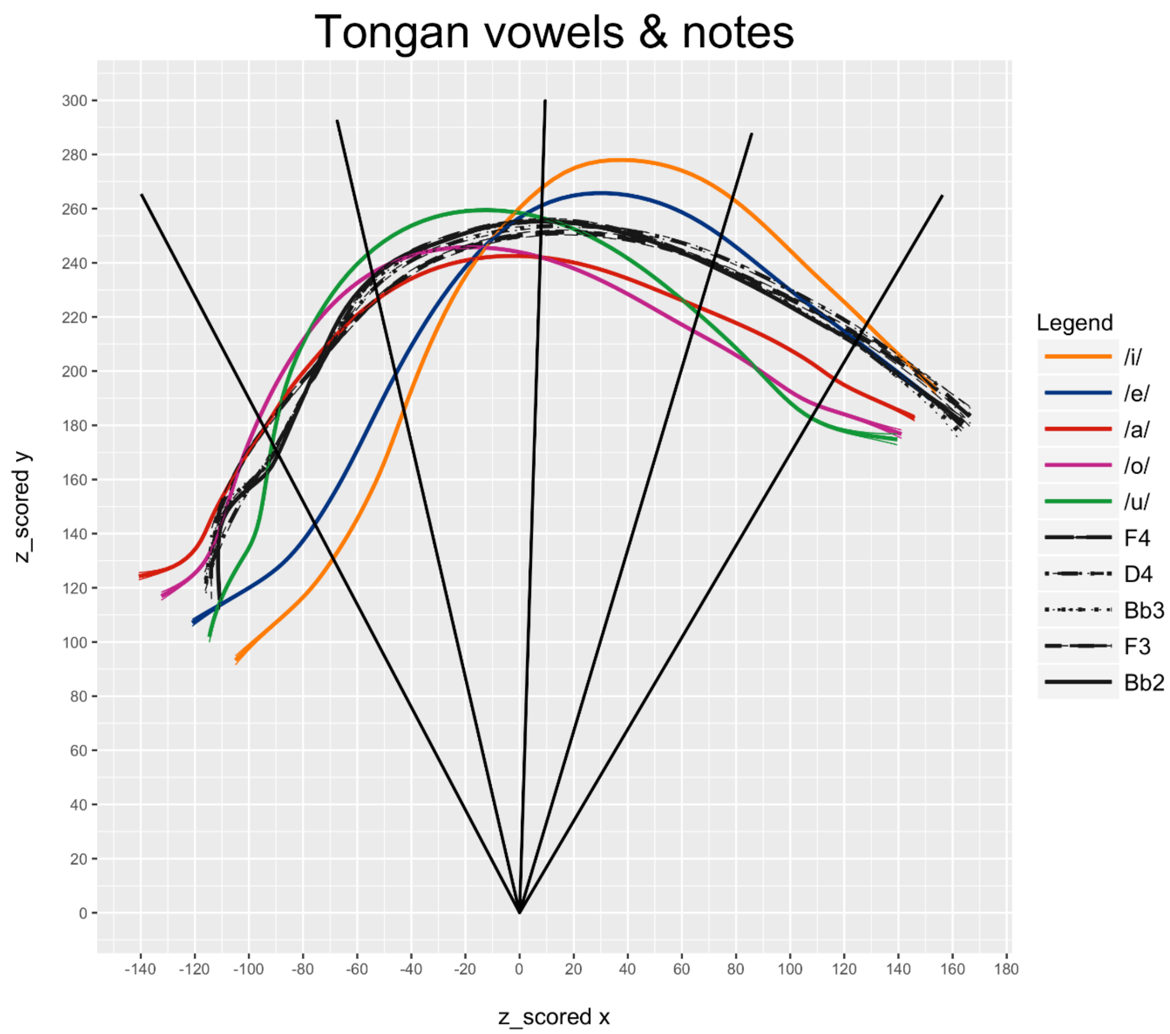


Figure 9.3: SSANOVA average curves for the five vowels of Tongan and five different sustained notes played on the trombone by the same speakers (otherwise identical to figure 9.1 above).

For the Tongan data shown in figure 9.3 above, it looks like the back of the tongue assumes a position during trombone playing that approximates the position employed when producing the back vowels /o/ and /u/ and the low vowel /a/ during speech. The front of the tongue, however, assumes a more elevated position than that used for the back and low vowels, matching the tongue height for the front vowels /e/ and /i/. Not much difference in tongue position can be observed between the average curves for the various sustained notes included in the figure, ranging from the low notes Bb2 and F3 through the middle range of the instrument (Bb3) to the high range (D4 and F4); due to their small differences, these notes will be grouped together as low notes (Bb2 and F3), middle notes (Bb3), and high notes (D4 and F4) in some of the following figures to enhance the readability of the plots.

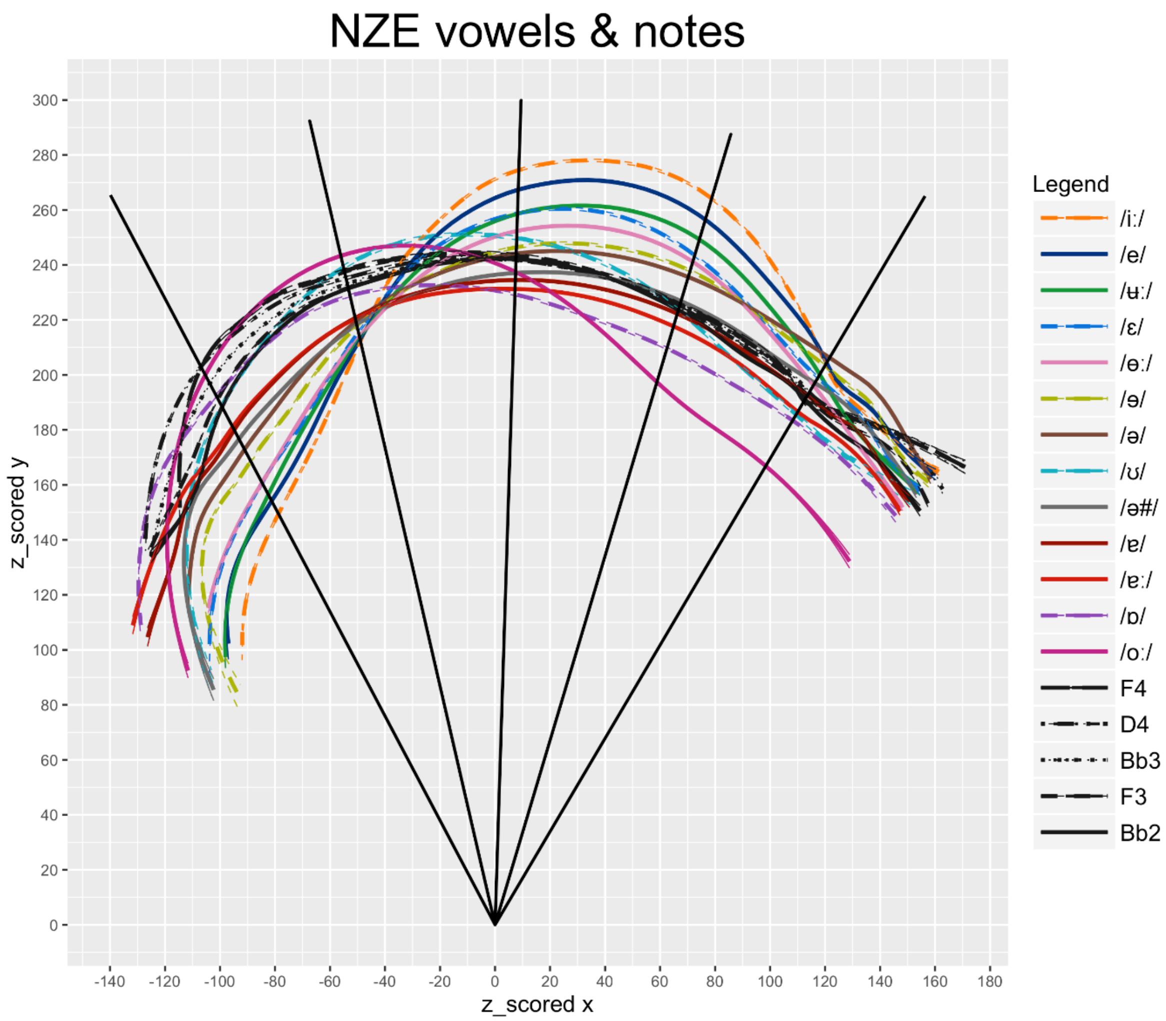


Figure 9.4: SSANOVA average curves for the monophthongs of NZE and five different sustained notes played on the trombone by the same speakers (otherwise identical to figure 9.2 above).

The combined plot of NZE vowels and sustained note average contours above (figure 9.4) reveals a situation similar to that described for the Tongan data above. As in the case of the Tongan participants, the average contours for the various sustained note productions seem to pattern with the back vowels when considering the back of the tongue; the front of the tongue, however, seems to assume a somewhat lower position than the one utilized by the Tongan participants, falling within the medium range of tongue front height among the NZE monophthongs.

9.3 Comparison of the two language groups

Since all the ultrasound data were z-scored to resemble a similar scan depth and probe angle in relation to the individual participants' oral cavities, it is possible to overlay the average tongue contours for sustained note production by one language group directly onto the average note productions by the other. This is shown in figure 9.5 on the next page, indicating that, overall, the position of the back of the tongue is more variable for the trombone players whose L1 is NZE, and that the back of their tongue seems to retract more during trombone playing than that of the Tongan trombone players. The NZE-speaking players also position their tongue lower than the Tongan players for the length of the tongue extending from its midsection to the front. Regarding the change in tongue height for lower versus higher sustained notes played on the trombone, there are two opposing patterns shared by participants across both languages. A majority of players (eleven of nineteen participants) utilize a more elevated tongue position for higher notes, with only five players showing the opposite pattern; three individuals hardly change their tongue position at all or display an inconsistent pattern throughout the range of the trombone. The majority pattern is reflected in the figures by the higher average curves for high notes displayed by both language groups, which result from z-scoring and averaging the data across all participants in each group. Whether a player elevates or lowers their tongue for high notes does not seem to correlate with playing proficiency or any other data collected from participants; all relevant participant information was listed in sections 7.7.1 and 7.7.2, and individual plots of all participants showing the average tongue contours for monophthongs and sustained note productions are included in appendix C, pages 265-284.

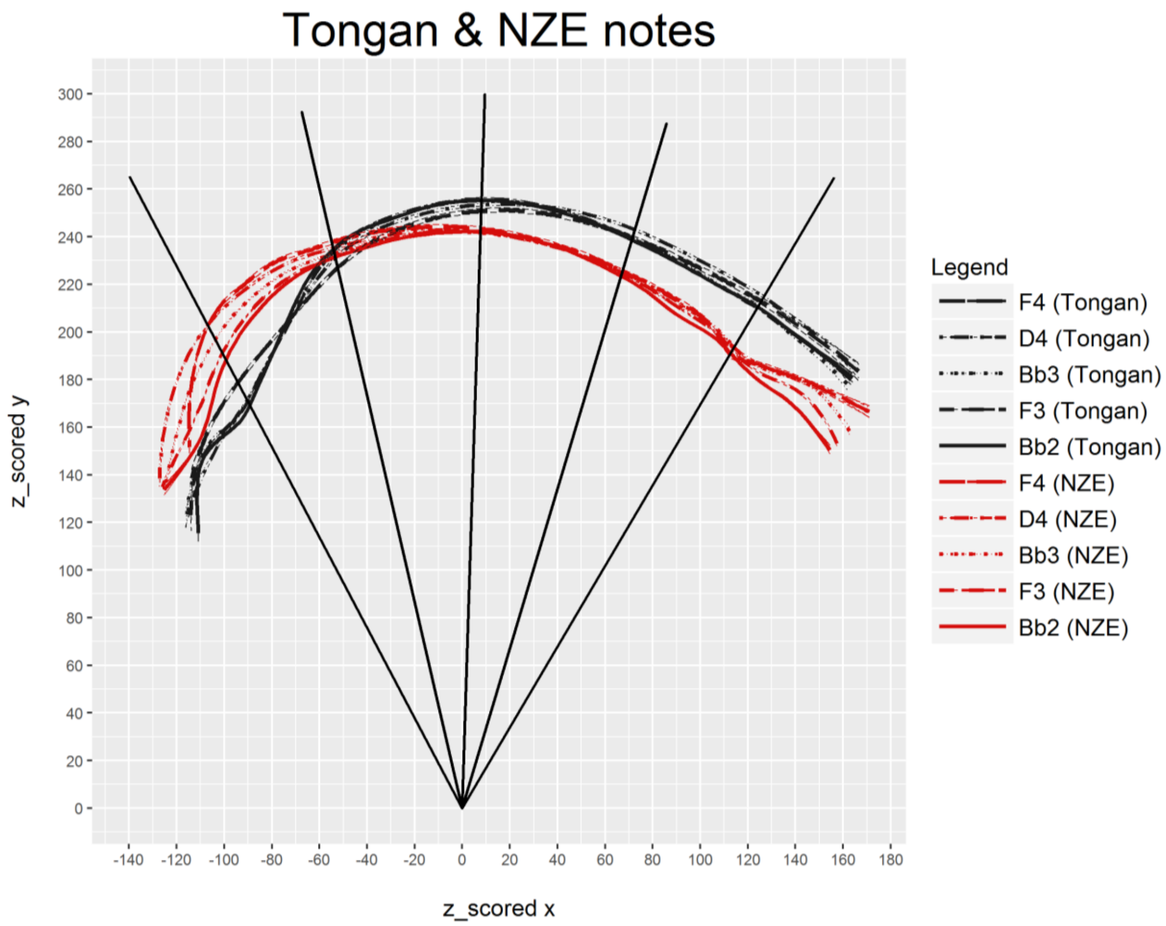


Figure 9.5: Average tongue curves for sustained note productions by trombone players from the two different language groups (black: Tongan, red: NZE).

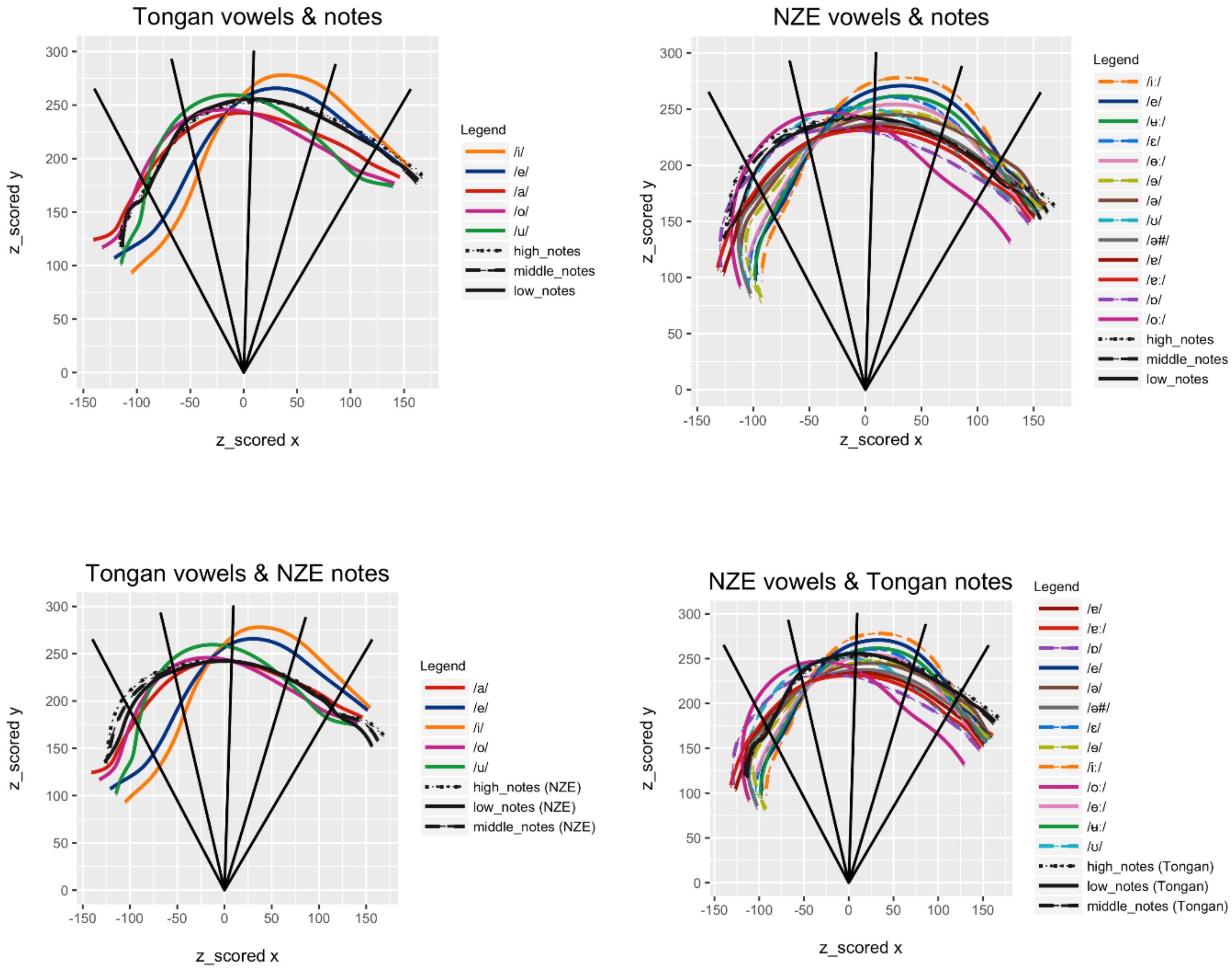


Figure 9.6: Matching (top row) and mismatched plots (bottom row) for the two language groups' vowel productions, overlaid with sustained note productions by speakers of the same (matching) or opposite (mismatched) language group.

A different approach to visualizing the differences among the two language groups is taken in figure 9.6 on the previous page, which overlays the average note contours for one language group onto the monophthong productions of the same, and opposite, groups. The resulting matched (top row) and mismatched (bottom row) plots allow the reader to directly observe what kind of differences in the averaged monophthong productions might lead to the playing differences illustrated in figure 9.5. Note that only three SSANOVAs curves are printed for the sustained notes, which have been conflated into three groups as described earlier on. It should be easy to observe that the sustained note average contours in the mismatch plots do not match up as well with the respective language's average vowel tongue positions as in the properly matched conditions. In the case of the sustained note averages produced by Tongan participants plotted with average NZE vowel articulations (bottom left plot), the positioning of the back of the tongue does not seem clearly mismatched; however, the position of the front of the tongue is higher than for any of the NZE monophthong average contours. For the NZE note productions overlaid on the averaged Tongan vowel contours (bottom right plot), the back of the tongue is positioned more posteriorly while the position of the front of the tongue seems to be a reasonable fit. Considered in context, these observations seem to suggest that a player's tongue position during sustained note production matches up with certain features of the tongue positions used during native vowel articulation, while this is not the case as much in the mismatched conditions; furthermore, these observations refer to different vowel tongue positions when considering the placement of two different sections of the tongue, the back and the front, during sustained note production.

While the average tongue positions during sustained trombone note production are clearly different for the two language groups as shown in figure 9.5 above, average tongue contours for a subset of monophthongs of both languages that can be expected to feature relatively similar articulations across the two languages (based on their acoustic descriptions), map up fairly well, as shown in figure 9.7 on the next page. Acoustic descriptions of NZE in the literature (cf. section 7.3.1) indicate that NZE DRESS (/e/) is close compared to a more 'cardinal' pronunciation of the /e/ vowel in Tongan; similarly, the NZE THOUGHT vowel (/oː/) is comparatively raised, possibly due to a chain shift documented for other varieties of English that motivates it to move into the space vacated by the fronted GOOSE vowel (/ʉː/) (Ferragne & Pellegrino, 2010, p. 30; cf. Scobbie, Stuart-Smith, & Lawson, 2012; Stuart-Smith et al., 2015).

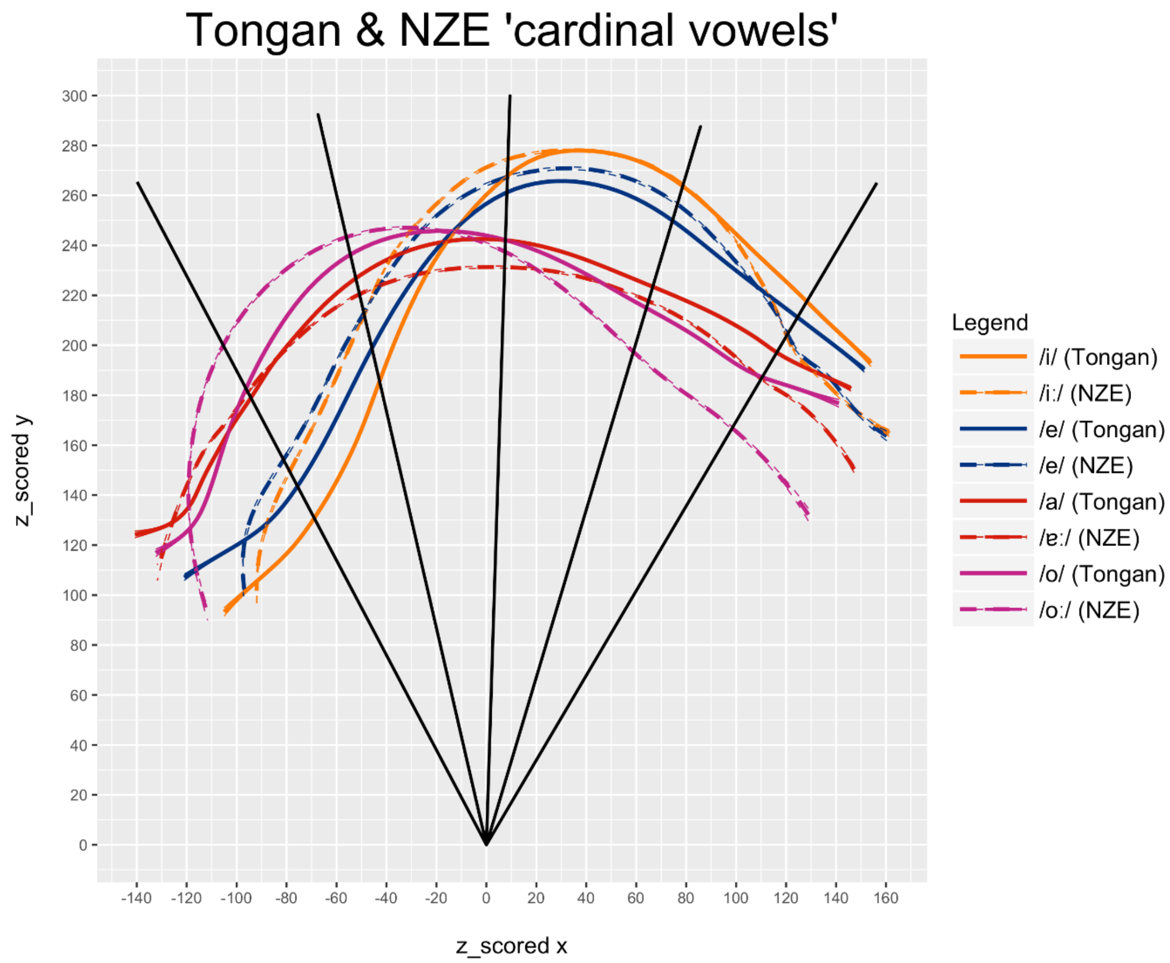


Figure 9.7: Average tongue contours for selected monophthong productions by participants from the two different language groups (Tongan = solid lines, NZE = dashed lines); only vowels were selected that can be expected to be roughly similar in terms of their acoustics across the two languages.

9.4 Quantification of differences between tongue contours for vowel and trombone note productions

While the statistics of SSANOVAs inform us that average curves are statistically different when their 95 percent confidence intervals (1.96*SE) do not overlap, the technique provides no way of determining whether the difference between two hypothetical curves A and B is smaller, the same, or bigger than the difference between the hypothetical curves A and C. To address this issue, I developed a measurement that calculates the absolute area between two average curves in polar coordinates, relative to the estimated ultrasound transducer position. This measurement was implemented in MATLAB (Mathworks Inc., 2015) as it requires interpolation in polar space; although such a function does not come with the standard

MATLAB license, a function called ‘interp.polar’ is available on github (Moerman, 2014). My script requires two Theta values (the first part of each polar coordinate specification, indicating the angle of the respective fan line extending from the origin) restricting the angle for which the measurement is going to be carried out, interpolates each of the two average curves to 1001 equidistant Theta values falling within the specified angle, and then sums up the absolute values of the area differences between the very pointed triangles connecting the origin with two consecutive interpolated points lying on each average curve. Note that the difference is calculated accurately even when two curves intersect within the specified measurement interval, due to independent comparisons of a large number of individual subsections (1000) and summing up their absolute values. Figure 9.8 below illustrates this procedure; note that only eight sections are shown in the zoomed in image, while in reality the angle was divided into one thousand such sections.

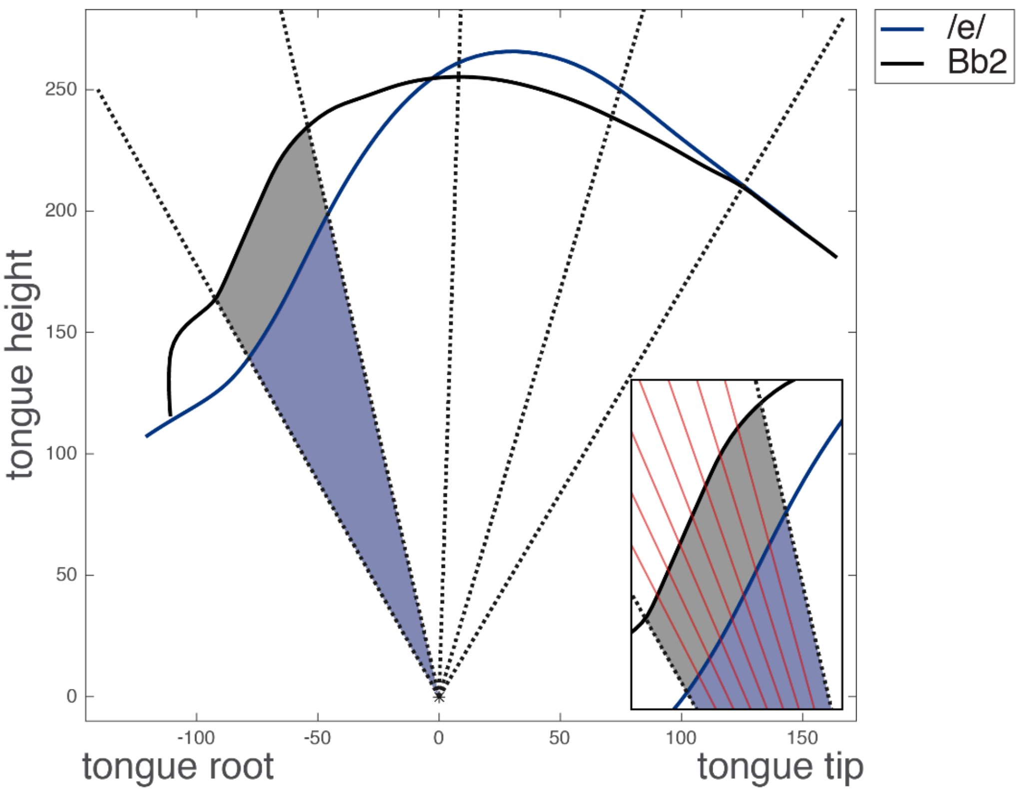


Figure 9.8: Schematic illustration of the area difference measurements implemented in MATLAB, forming the basis for the numerical comparisons in the following figures.

Since the results presented above indicated different patterns for the back and front of the tongue for both of the two language groups, I decided to divide the average tongue contours into four different sections, shown by the fan lines superimposed on all previous plots. This division was motivated by findings demonstrating the functional independence of various sections of the tongue (cf. section 3.1.3). While the inconsistency of SSANOVA curves at the edges determined the placement of the extreme fan lines, the decision to further divide the fan-shaped area into four equally sized sections reflects the methodological considerations in Stone, Epstein and Iskarous' (2004) paper. Nonetheless, the match is not a perfect one as the use of a small transducer seems to have allowed me to capture a larger section of the tongue than the setup by Stone, Epstein and Iskarous. Figure 9.9 below shows the average tongue contours for the five vowels produced by the Tongan participants (pooled, z-scored data) overlaid onto a figure from Stone, Epstein and Iskarous' publication.

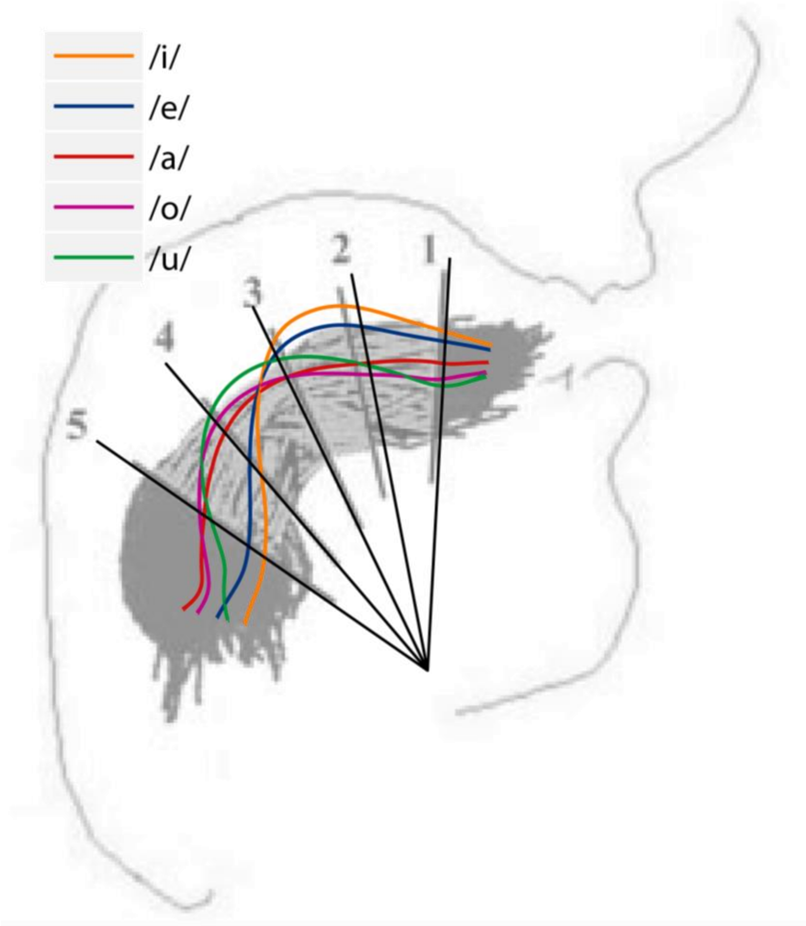


Figure 9.9: Average tongue contours for the five vowels of Tongan, produced by, and normalized across the ten Tongan participants of this study, overlaid on an image taken from figure 3: Five tongue segments of the Fisher-Logemann dataset; reproduced with permission from Stone, Epstein and Iskarous, 2004, p. 511.

Using the area differences calculated in MATLAB, I created the line plots shown in figures 9.10a and b on the following pages, which report the area difference values for the back (a) and front of the tongue (b); the arrangement of the figures corresponds to the match/mismatch plots in figure 9.6 above. The order of the vowels on the x-axis was chosen to reflect the relative 'backness' (figure 9.10a) and 'frontness' (figure 9.10b) for the various monophthongs in the respective languages. The plots in the top row of figure 9.10a nicely illustrate that the position of the back of the tongue for sustained note productions patterns with the back vowels in each language, while the mismatched data (bottom row) show a more complicated pattern. For the front of the tongue (figure 9.10b), the data do not seem to show a consistent pattern at all, suggesting that the position of the front of the tongue during sustained note production is not related to vowel production in a player's native language. Note, however, that there seems to be a clear mismatch at the very front of tongue (past the area selected to carry out numerical comparisons) for sustained note productions by Tongan players overlaid onto the NZE vowel system (figure 9.6 bottom right plot, average curve shape to the right of the rightmost fan line); this area, however, had to be eliminated from the nurmerical comparisons mostly due to irregularities in average curve shape for sustained note productions by the NZE participants.

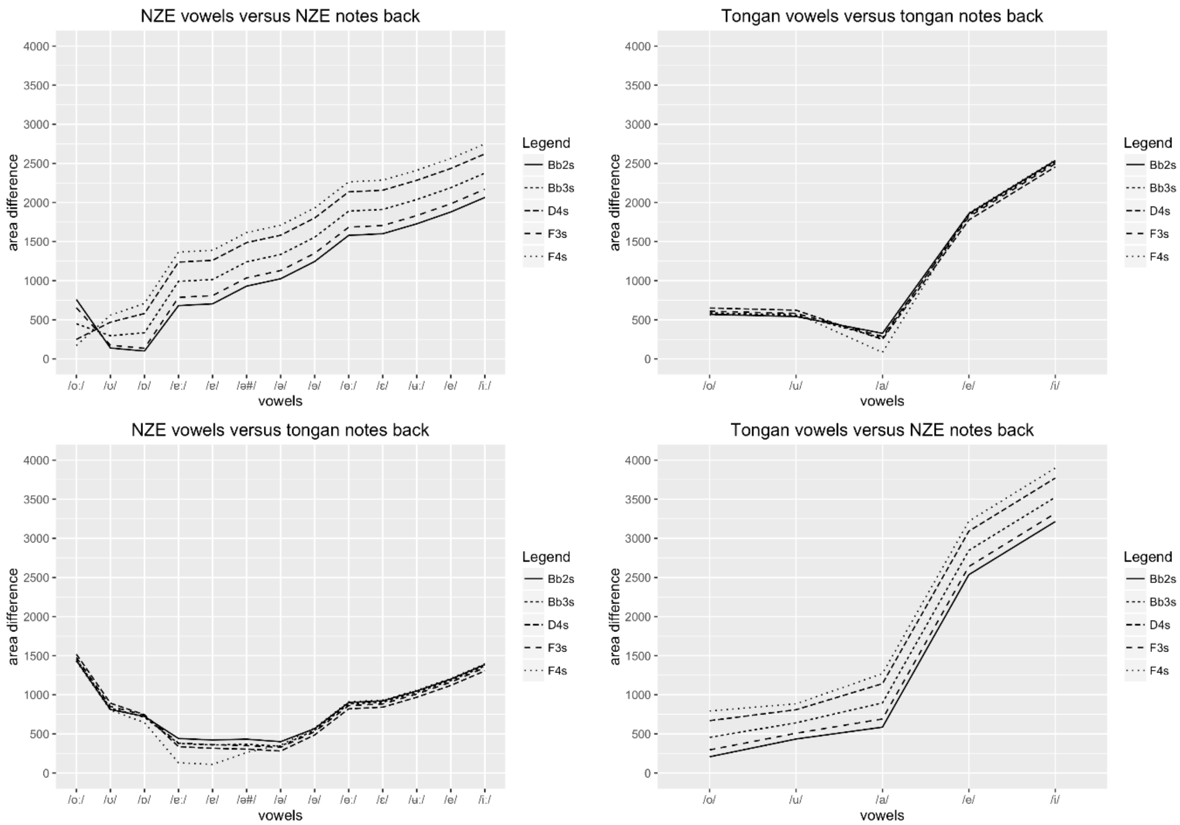


Figure 9.10a: Line plots reporting the measured area in between average tongue contours for vowels (x-axis) and the five different notes (y-axis) for the back of the tongue (wedge-shaped area between two leftmost fan lines overlaid on all preceding figures in this section). The ordering of the different vowels corresponds to the order of ‘backness’ in the respective languages; note that the scale is arbitrary due to the underlying z-scored data. The overall arrangement of the figure corresponds to figure 9.6 above.

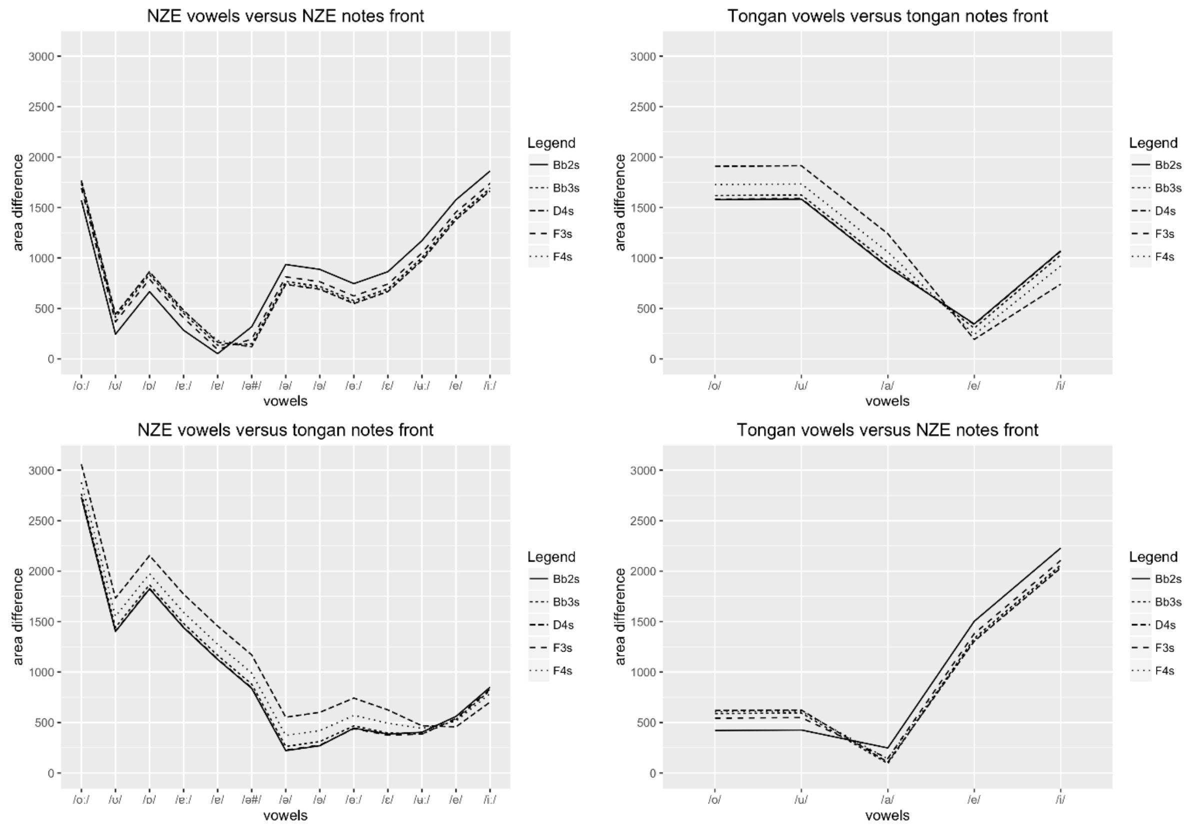


Figure 9.10b: Line plots reporting the measured area in between average tongue contours for vowels (x-axis) and the five different notes (y-axis) for the front of the tongue (wedge-shaped area between two rightmost fan lines overlaid on all preceding figures in this section). The ordering of the different vowels corresponds to the order of vowel 'frontness' in the respective languages; note that the scale is arbitrary due to the underlying z-scored data. The overall arrangement of the figure corresponds to figure 9.6 above.

9.5 Results for individual participants

While the pooled data reported above are representative of the overall distributions for the two language groups, they hide some interesting patterns pertaining to smaller groups of individuals within and across the two language groups. One of these is the observation that the average tongue contours for sustained note production on the trombone seem to pattern on a selected vowel tongue position for most of the length of the tongue for a subset of participants; at the same time, this is not the case for the majority of individuals. Specifically, the tongue position during trombone playing seems to pattern with a central vowel tongue position for four NZE-speaking participants, and with a high back vowel for two Tongan participants; an earlier version of my hypothesis regarding language influence on brass playing (see Heyne & Derrick, 2014; Heyne & Derrick, 2015a&b) was informed by this observation as a majority of early participants seemed to fit this pattern.

9.5.1 NZE players who use a centralized tongue position during sustained note production

Figure 9.11 on the next page shows the average tongue contours during monophthong and sustained note production (grouped as high, middle and low notes) for four NZE participants who employ a centralized vowel tongue position during playing, out of a total of nine NZE-speaking participants in this study. The plots are arranged in the order of how well the individual participants' results fit with the pattern described above. S1 was the pilot participant for this study and the SSANOVA curves are thus based on a much smaller number of tokens than for the other participants (see individual token counts provided in section 8.7.3); his average tongue contour for low notes very closely approximates the average vowel contours for non-final (/ə/) and final schwa (/ə#/), although the very back and front of the tongue are positioned higher. He also displays large differences for notes falling into different registers at the back of the tongue. S12 also patterns with the average contours for non-final and final schwa, however, the front of his tongue trails off toward the START vowel (/ɐː/), and changes in tongue position for the different registers are much smaller than for S1. S5 overall positions the tongue lower than even for non-final and final schwa, towards the STRUT vowel (/ɐ/), but positions the back of the tongue higher than the other participants included in the figure. This participant shows very little change for playing in different registers; note, however, that the ultrasound probe orientation is different from most other participants, leading to a loss of data at the back of the tongue. Finally, S26 patterns similarly to the other players in this group, however, it is his middle and high note productions that approach the vowel averages for non-final and final schwa, while the position used for low notes is much higher. Interestingly, all individuals in this group are very proficient players of the trombone (three professionals: S1, S12, and S26, and one semi-professional: S5).

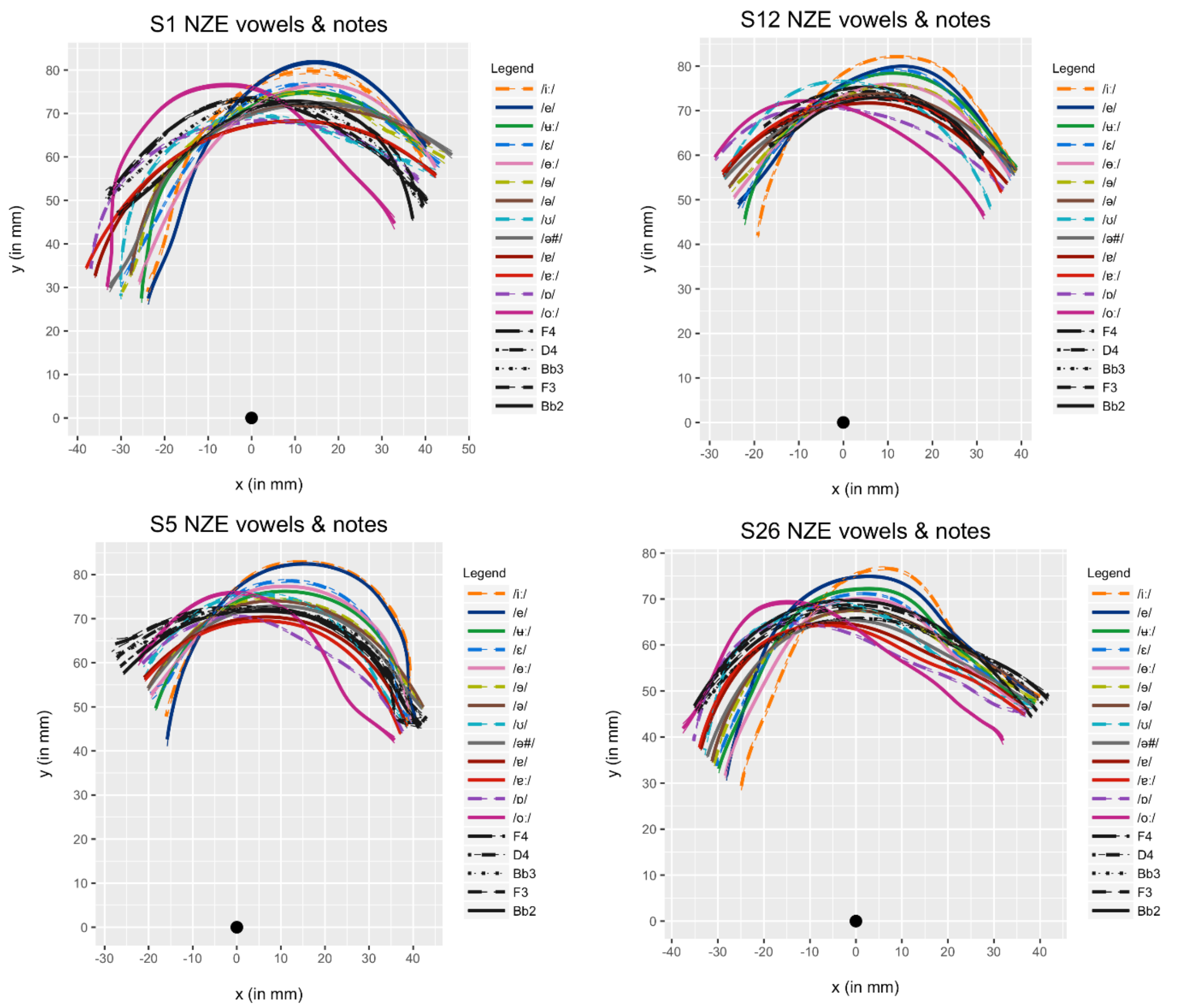


Figure 9.11: Plots for 4 NZE participants whose average midsagittal tongue curves for sustained note production on the trombone pattern closely with or within the vicinity of one of the centralized vowels of NZE. The scale for these plots is in mm and reflects the different sizes of the players' oral cavities (these data have not been rotated or normalized/z-scored).

9.5.2 Tongan players who use a back vowel tongue position during sustained note production

In analogy to figure 9.11, figure 9.12 on the following page illustrates a similar situation for some of the Tongan players, whereby a large section of the average tongue contours for sustained note productions pattern with one of the back vowels in the language; in the final dataset, only two out of ten total Tongan subjects display such a pattern. For S4, the average contour for high notes clearly patterns with the back vowel /o/, while notes in the lower registers display a more elevated back of the tongue. For

S14 the pattern is similarly strong but follows the high back vowel /u/ while the front of the tongue is held much higher than for /u/.

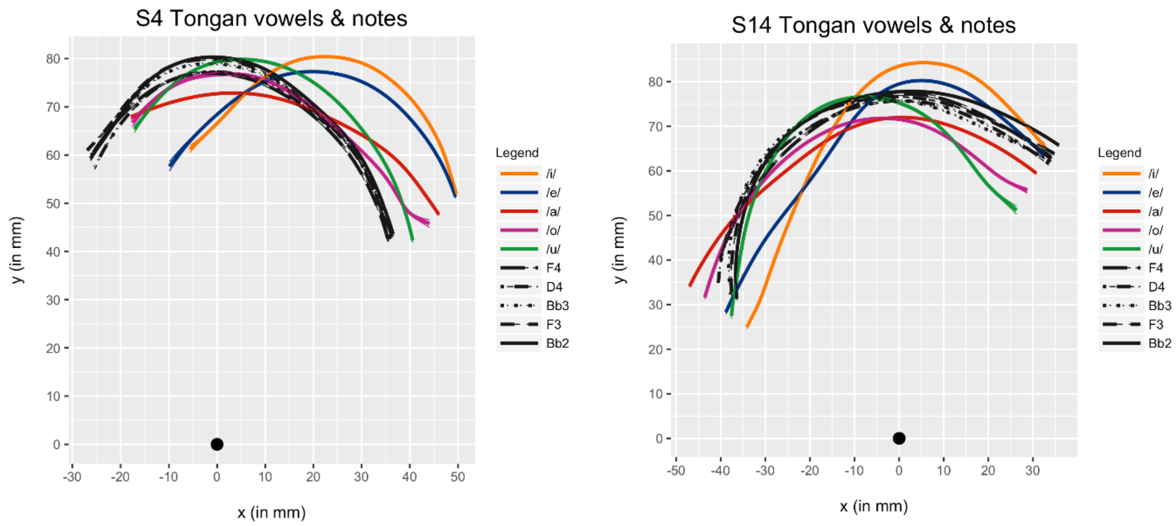


Figure 9.12: Plots for two Tongan participants whose average midsagittal tongue curves for sustained note production on the trombone pattern closely with or within the vicinity of one of the back vowels of Tongan. The scale for these plots is in mm and reflects the different sizes of the players' oral cavities (these data have not been rotated or normalized/z-scored).

9.5.3 Participants displaying a pattern more typical for the opposite language group

A small number of players in both groups seem to adhere to the pattern more typically displayed by players from the opposite language group. Figure 9.13 on the next page features the plots for two players each from the NZE and Tongan language groups. Overall, the patterns are not as clear as in the two previous sections, and of course, the Tongan language does not have any central vowels, meaning that a central position represents a fictional reference for the Tongan participants included here. Individual plots including separate SSANOVAs for the five different notes for all participants can be found in appendix C, pages 265-284.

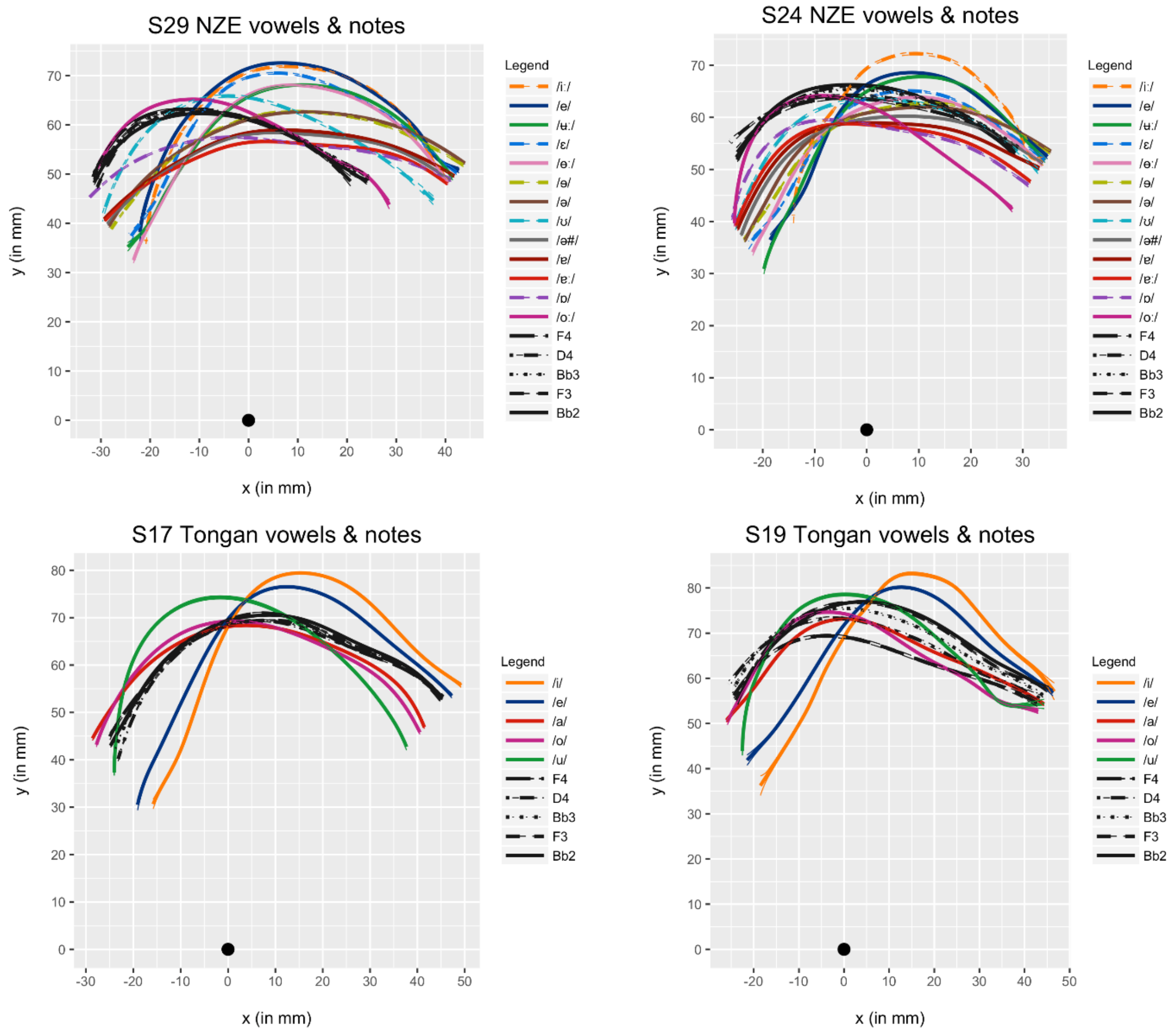


Figure 9.13: Plots for two NZE and two Tongan participants whose average midsagittal tongue curves for sustained note production on the trombone assume a position that fits more closely with the pattern more typically displayed by players from the opposite language group. The scale for these plots is in mm and reflects the different sizes of the players' oral cavities (these data have not been rotated or normalized/z-scored).

9.5.4 Consonant production and trombone articulation data for two participants

While I did not have enough time to analyze the place of articulation for selected consonants and the place of attack during trombone playing for all participants of this study, I managed to complete such analyses for two early participants. These individuals are S4, the amateur Tongan player recorded in Christchurch, and S5, a semi-professional participant whose L1 is NZE.

For both participants, I calculated SSANOVA average curves for the coronal stops used in each language (dental /t/ only for Tongan; NZE has alveolar /t/ and /d/) by averaging the traces from individual ultrasound frames judged to represent the maximum constriction falling within each phoneme's duration, based on the speech audio segmented manually (cf. section 8.3.2). A similar approach was taken for the coronal articulations produced during brass playing by first eliminating all notes shorter than eighth-notes (to exclude possible velar attacks that can occur during double-tonguing), and subsequently identifying the appropriate ultrasound frame to trace which displayed the maximum constriction at the beginning of each note. Furthermore, I identified tongue steady states assumed during pauses from speaking and trombone playing to calculate lingual averages for the respective interspeech and inter-playing postures (cf. sections 3.1.2 and 3.6.1). ISPs resulted naturally from the timing of the slide transitions set up in Powerpoint, while I specifically added sufficiently long rests to the musical passages to elicit IPPs. The relevant ultrasound frames were selected manually and care was taken not to select any frames where participants were bracing their tongue against the roof of the mouth, swallowing, or inhaling. All calculations were performed in the same manner as for the vowels (see sections 8.6-8.8); however, traces were only cut at the edges where less than ten percent (twenty percent was used for vowels and notes) of combined traces provided more extreme points (in terms of their Theta or angle values).

The left panels of figures 9.14 and 9.15 on the next page show SSANOVA average curves for initial coronal consonants produced during wordlist reading and interspeech postures (ISPs) overlaid on the respective monophthong productions (faint lines). The right panel of each figure again plots the same average curves for stop consonants and ISPs (faint lines) but highlights the relationship of coronal articulations produced during trombone playing, inter-playing posture (IPP), and the tongue position assumed during sustained note production; table 9.1 lists the number of individual tokens used to calculate these SSANOVA average curves.

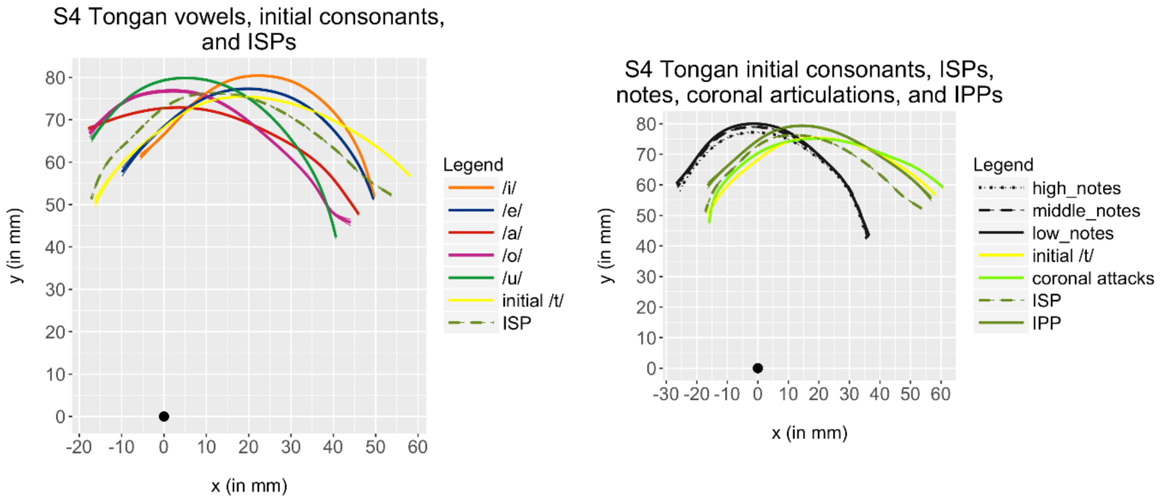


Figure 9.14: Average tongue contours for initial dental /t/s and interspeech postures (ISPs) produced by S4 Tongan, overlaid on monophthong productions by the same participant (faint lines; left panel). Comparison of average tongue positions assumed during sustained note production in different registers of the trombone, coronal articulations produced during trombone playing and inter-playing postures (IPPs), overlaid on initial consonant productions and ISPs by the same participant (right panel).

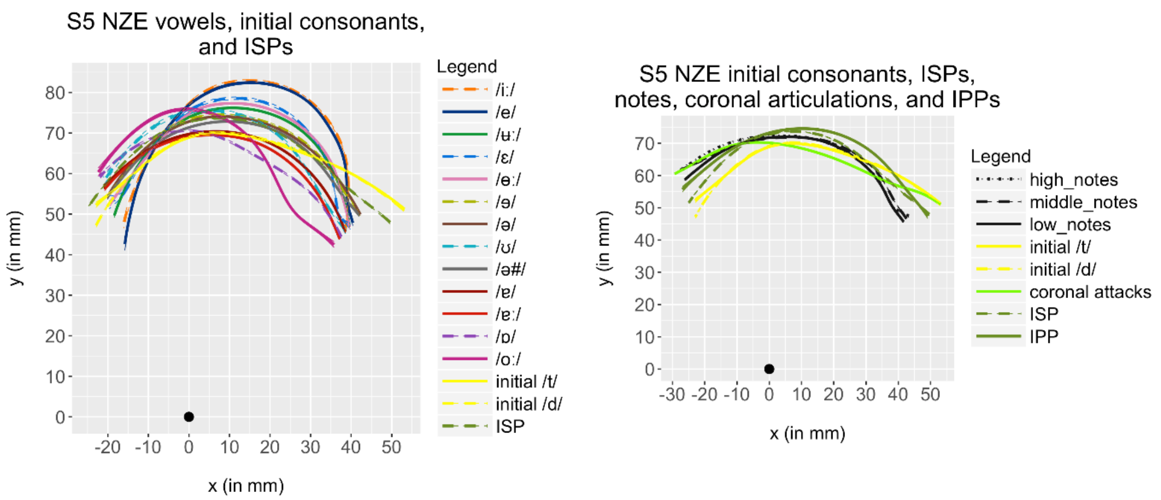


Figure 9.15: Average tongue contours for initial alveolar /t/s and /d/s and interspeech postures (ISPs) produced by S5 NZE, overlaid on monophthong productions by the same participant (faint lines; left panel). Comparison of average tongue positions assumed during sustained note production in different registers of the trombone, coronal articulations produced during trombone playing and inter-playing postures (IPPs), overlaid on initial consonant productions and ISPs by the same participant (right panel).

	playing proficiency	initial coronal consonants		ISP	Coronal articulations during trombone playing	IPP
		dental /t/ (S4) alveolar /t/ (S5)	**alveolar /d/**			
S4 Tongan	amateur	64	N/A	138	346	13
S5 NZE	semi-professional	108	88	173	539	37

Table 9.1: Token numbers for initial coronal consonants, ISPs, coronal articulations, and IPPs produced by S4 and S5.

Recall that ISPs, and similarly, IPPs, can be assumed to arise out of motor efficiency requirements relating to token or type frequency of articulatory targets. The fact that brass players almost exclusively use coronal attacks (cf. section 3.6.1) sets up expectations of a clear connection between IPP, the place of articulation used for initiating notes, and the average tongue shape assumed during sustained note production; both participants included here seem to fulfill these expectations, with the Tongan player having to cover a much larger physical distance between the place of articulation and the steady-state assumed during note production. It is also striking how much the tongue shapes for coronal consonant productions differ between the speakers of the two languages. Presumably, the dental articulation of /t/ by the Tongan speaker requires the tongue to be brought much further forward, explaining the overall much higher and more fronted tongue position compared to the NZE speaker.

9.6 Chapter summary

Chapter 9 of this thesis presented pooled and individual results for the nineteen participants observed in this thesis, divided into two language groups. A clear difference in the tongue position used during sustained note production on the trombone was shown on the language group level by overlaying the average tongue shapes for averaged note productions of both groups in a single plot. Average tongue contours during sustained note production were also plotted with average vowel tongue contours for either language, illustrating that they do not map up for the whole length of the tongue. Separate comparisons for the back and front of the tongue, however, indicate that the position of the back of tongue during note production patterns with the back vowels for each language, while a vowel-unrelated constraint seems to lead to a difference at the front of the tongue. At the end of this chapter, I also presented limited place of articulation data for one player from each language group, showing the average tongue contours for coronal consonants produced during speech, and coronal attacks produced during trombone playing, as well as inter-speech (ISPs) and inter-playing postures (IPPs).

10 Discussion

In this chapter, I will re-examine the results presented in the previous sections in light of the hypotheses formulated in chapter 4, and relate them to various issues arising from the existing literature. Specifically, separate constraints arising from airflow and acoustical requirements, and considerations of articulatory efficiency, will be elaborated, and it is argued that they represent higher-order effects mediating language influence on brass playing. I will also provide a short discussion in which I argue that these findings are the result of language influence rather national schools or styles of playing, and consider the implications of my findings for modular theories of motor control and brass playing.

10.1 Evaluation of hypotheses

Hypothesis 2a) (cf. section 4.3) postulated that tongue positions during sustained note production on brass instruments would be based on motor memory, and that such motor memory would be the result of speech articulation, specifically the tongue shape of vowels. The pooled results presented in the previous chapter based on the z-scored data (sections 9.3 and 9.4) do not provide support for such a pattern holding for the entire tongue shape. Separate comparisons for the back and front of the tongue, however, show that the position of the back of the tongue during sustained notes played on the trombone patterns with the back vowels for each language group, providing support for hypothesis 2b). Furthermore, a vowel-unrelated constraint seems to lead to a pronounced difference at the front of the tongue.

Specifically, the data show that Tongan players use an overall higher tongue position that patterns with the back vowels (and the low central vowel /a/) in the Tongan language for the back of the tongue (tongue root), while the front of the tongue assumes a position comparable to that used for the front vowel /e/ (and /i/ for high notes). NZE trombone players similarly use an average tongue contour during sustained note production that patterns with the back vowels of their language for the back of the tongue (tongue root); the front of the tongue, however, assumes a position comparatively lower than that used by the Tongan players, at medium height compared to monophthong productions in NZE (cf. figure 9.5).

The results presented in this thesis thus support hypothesis 2 through its extension 2b: Tongue positions assumed during sustained note production on brass instruments are based on motor memory, and functionally independent sections of the tongue are individually affected by motor memory from a player's native language, in agreement with a modular theory of motor control.

A discussion on how these findings fit with a modular theory of motor control will be presented below in section 10.11.1. Recall also that hypothesis 1 - Brass players can perceive (consciously or subconsciously) the acoustic consequences of playing differences between players with different native languages - was cautiously supported by the questionnaire data presented in chapter 5, which showed that brass players at least believe they can perceive differences in the playing by performers with different native langauges.

10.2 Group-level findings versus individual variability

Plots of individual participants' average monophthong and sustained note productions (section 9.5, appendix C, pages 265-284) showed considerable individual variability, in agreement with earlier research documenting tongue positions during brass playing (cf. section 2.5 above). Furthermore, even though the SSANOVAs plotted in the previous section feature very small error bars due to averaging across large numbers of observations (except for the pilot participant), both monophthong and sustained note productions display substantial variability at the individual token level. The question, then, is how the underlying variability leads to the overall difference observed at the group level.

I propose the following explanation: Language-specific motor memory for vowel tongue positions imposes an additional constraint on the tongue positions that trombone players are likely to employ during sustained note production, secondary to more basic constraints arising from airflow and acoustical requirements, and considerations of articulatory efficiency. The functional independence of various sections of the tongue and the modular organization of the vocal tract musculature (cf. sections 3.1.2, 3.2.3 and 3.3.3) additionally mean that the various constraints can act upon individual sections or on the tongue as a whole. Finally, the average tongue position used by individual players during sustained note production on the trombone is understood to reflect a local optimum rather than a globally optimized vocal tract configuration mediated by all relevant factors. All of these considerations are discussed separately in the subsequent sections, followed by a general discussion considering the various points in combination.

10.3 Tongue position during brass playing as local optimization

In line with the findings by Ganesh et al. (2010) and the theoretical argument laid out in Loeb (2012; cf. section 3.2.4), this thesis takes the view that the tongue position assumed during sustained note production on the trombone represents a local solution to a complex optimization problem with many ‘good-enough’ solutions (Loeb, 2012). Ganesh et al. (2010) assert that the CNS determines the appropriate motor behavior in a “tradeoff of motor memory, error, and effort minimization” (p. 382). With motor memory in our case presumably arising from native speech articulation, what are the other factors governing this tradeoff?

For a beginning player, error might simply equate to the rate of unsuccessful attempts at producing a desired note; however, for an advanced player, error might be evaluated in regard of the aesthetic quality of the produced sound, with perceptual abilities likely to evolve with increasing proficiency (cf. section 2.6 above). Effort minimization could be applied directly to the physical position of the tongue by minimizing the distance to be covered between the coronal place of articulation and the sustained note position (cf. section 9.5.4). It could also take place at other locations within the sound-producing mechanism, however, including, for example, the amount of airflow required to produce a certain note while utilizing a specific tongue position.

Accepting that the control of vocal tract movements during brass playing is governed by local optimization suggests an explanation for the astonishing amounts of individual variability observed in this and earlier empirical studies on brass playing. It is conceivable that beginning players might initially explore different local optima (different vowel tongue positions but possibly also completely new, language-unrelated gestures) before settling on a more stable default position that would be locally optimized using acoustic/auditory information (cf. section 3.3) and effort minimization. Employing a configuration frequently executed in speech (such as a vowel tongue shape) would seem to reduce both error and the required effort, at least until a player has developed sufficient motor memory for the new motor action. This framework would also allow for players to gradually ‘unlearn’ (cf. Heyne & Derrick, 2015b, p. 7) their language-related tongue shapes by developing brass playing-specific motor memory.

10.4 Requirements of airflow during brass playing

10.4.1 Overall midsagittal tongue shape

Trombone playing requires a greater amount of airflow than speech production, necessitating brass players to keep their vocal tract relatively unconstrained all the way from the lungs to the embouchure (where the airstream enters the instrument). Early MRI research on speech production (Baer et al., 1991) has shown that for extreme vowels (high front and low back vowels), the airway is heavily constricted in the oral or pharyngeal cavities, respectively. This would reduce airflow and might explain why, on average, both Tongan and NZE participants model the position assumed during sustained note production at the back of the tongue on one of the back vowels in the respective languages.

A more straightforward expectation following from the above reasoning, however, would be that NZE players should use a central vowel tongue position during playing (Tongan players lacking such a position in their vowel system), as schwa has long been regarded as effecting the least constriction in the vocal tract (Fant, 1960; Johnson, 2011; Silverman, 2011; to name but a few). Indeed, this is what I found for a subset of the NZE participants (cf. section 9.5.1 and early hypotheses in Heyne & Derrick, 2014; Heyne & Derrick, 2015b); but why do not all NZE players utilize this position then? One clue could be that schwa is not as unconstrained as described in the literature; Gick (2002) used early x-ray data to show that at least for the four AE speakers in his data set, the pharyngeal cavity is significantly more constricted for schwa than for a lingual rest position assumed between utterances. The study also showed a difference for lexical versus functional schwa for one of the speakers (with functional schwa in the word ‘the’ being less constrained; cf. also Yamane-Tanaka, Gick, & Bird, 2004), which could mean that at least some of the NZE speakers recorded for this thesis might have a third separate articulatory target for schwa, in addition to lexical schwa occurring in non-final and final position. The different lexical schwas reported here clearly are not targetless, as established by their bimodal distribution and sociophonetic variability (cf. section 9.1 above). Unfortunately, no tokens of functional schwa are available to test this assumption for this study since all NZE speech data was elicited using word lists.

Returning to the Tongan players, positioning the back of the tongue in a location similar to the one used to articulate the back vowels /o/ (and in some cases /u/) might be the optimal solution for these players in terms of the aero-dynamical constraints outlined above. Based on the assumption that the pharyngeal constriction for Tongan vowels is comparable to the data for English reported above, the vocal tract configurations of the high vowels /i/ and /u/ might be too constrained in the oral cavity for sound production on the trombone, while /a/ might be too constrained in the pharyngeal cavity.

An alternative way of regarding the articulatory correlates of vowel tongue position is provided by John Esling's 'laryngeal articulator model' (2005), which represents an attempt at re-conceptualizing the traditional vowel quadrilateral based on "a growing body of articulatory evidence on pharyngeal phonetics" (p. 13). The model emphasizes the important (and neglected) role of "laryngeal articulator activity, controlling the pharyngeal resonator," and which is particularly relevant for the vowels classified as 'back' in the traditional representation (pp. 13-14). Figure 10.1 below presents Esling's revised vowel chart based on the 1996 IPA chart; note that Esling comments that "[t]he intersection of the three lines dividing the three regions in [the figure] should perhaps fall exactly on the location of schwa to represent the focal point of movement away from neutral toward any of the three directions," but that it was placed differently "to show the susceptibility of [ɐ] to becoming either front or retracted depending on the choice of articulator movement" (p. 23).

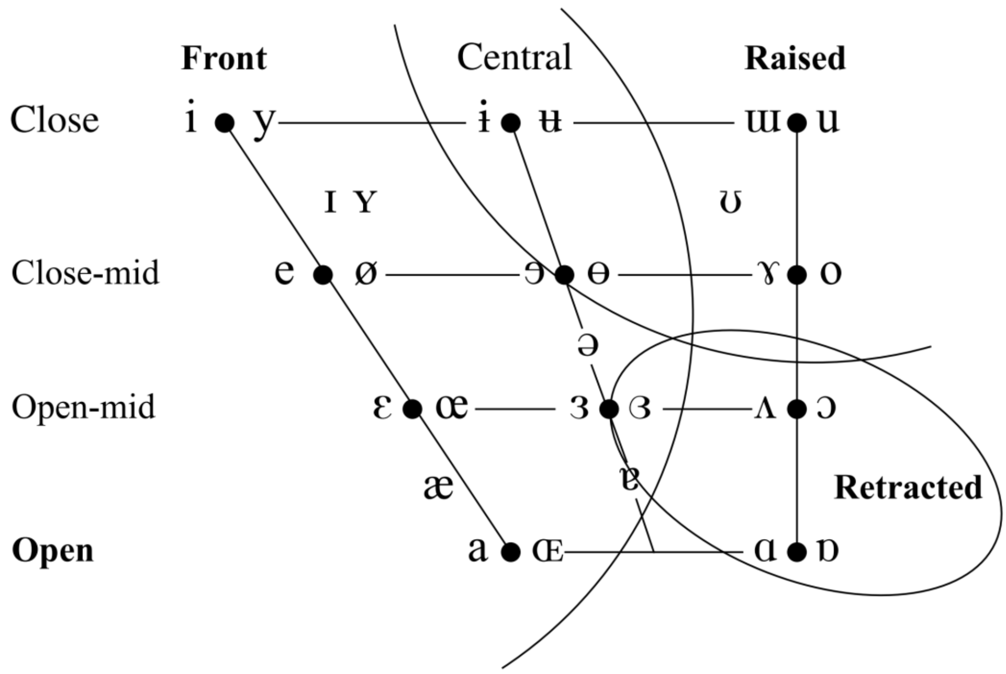


Figure 10.1: Revised vowel chart showing the division and overlap of articulatory regions. Reproduced with permission from Esling, 2005, p. 23.

Of interest to this thesis is the fact that the sustained note playing average contours for almost all participants across the two language groups fall within the area designated as ‘raised’ by Esling (optionally also including schwa, according to the comment above). Raising in Esling’s terminology refers to “the positioning of the tongue when it is high (pulled upward and backward),” in contrast to retracted vowels, for which “the lingual component” represents a “response to the sphinctering mechanism that closes the larynx” (p. 14); this action has certain consequences for the pharyngeal cavity that would seem to be advantageous concerning airflow and also acoustical considerations affecting vocal tract resonances. Quite recently, Esling, Moisik and Crevier-Buchman (2016) were able to present MRI data from two subjects that confirm some of the assumptions of the laryngeal articulator model. Specifically, the proposal that vowels pattern as front, raised and retracted in terms of larynx height was supported by the findings (p. 14). The close-mid back vowels (/ɤ, o/ in figure 10.1), however, were found to pattern in between raised and retracted (p. 14), precluding any clear implications for the tongue positions assumed during brass playing. A possible clue why a high raised (or close back) vowel tongue position might be unfavorable to brass playing is provided by the extreme values of larynx lowering (for modal voice) reported for both subjects of this study (pp. 12-13). Such a low position contrasts with larynx height measurements in a study on laryngeal activity during wind instrument playing by Rydell et al. (1996); these researchers found that, in contrast to flute and saxophone players, trumpet performers did not lower their larynx when playing, a fact they attribute to the high pressure used during trumpet playing (p. 47). I will return to considerations involving the larynx in section 10.4.3 below.

10.4.2 Air channeling within the oral cavity

Another aero-dynamical consideration concerns the channeling of the air inside the oral cavity. A renowned expert on the pathologies of wind instrument playing and tonguing, in particular, Wolfgang Angerstein from the university hospital in Düsseldorf, Germany, reports that the default is to form a median groove channeling the air from the larynx to the front teeth; if this channel trails off to the side, the air might miss the orifice of the instrument, leading to problems of articulation (Angerstein uses both midsagittal and coronal ultrasound for diagnosis). Interestingly, he reports that this problem may also show up in the speech of affected individuals as a lisp, suggesting that transfer does occur across these activities; however, such tongue movement disorders are overall quite rare (Härtel, 2012, p. 22; translation by myself). While the existence of (extensive) tongue grooving during wind instrument playing might seem to question the overall assumption that tongue shape during brass playing is modeled on vowel tongue positions, a finding from a pilot study by Lulich, Charles and Lulich (2016) employing three-dimensional ultrasound (and thus imaging in various sagittal and coronal planes) found that the tongue shape during clarinet playing matches the *shape* of the "ee" and "ah" vowels advocated in clarinet pedagogy but not tongue *position* (italics added). Even though I was not able to procure any information aside from the abstract for their presentation, I interpret the authors' wording to indicate that the three-dimensional shape of the tongue surface for clarinet playing resembles that used while articulating the vowel /i/ (the single subject of the study is a speaker of American English), although overall the tongue might be more retracted or fronted and lowered (or less likely, raised) during clarinet playing.

On the other hand, there is substantial evidence showing that tongue grooving also happens during speech production (Stone et al., 1988; Stone & Lundberg, 1996; Bressmann et al., 2005), albeit maybe less so than during wind instrument playing. The channeling of airflow within the oral cavity during both wind instrument playing and speech may be supported by findings documenting extensive lateral bracing as discussed in section 3.1.3 above. In their forthcoming paper, Gick et al. (in press) explicitly mention the role played by lateral bracing in creating "a consistent lateral seal along the tongue edge … basic to the production of any medial speech sound," which helps create the "closed aeroacoustic tube that directs the airstream through any medial speech constriction" (Gick et al., in press, p. 4). In fact, the articulatory stability gained from bracing against the rear molars might be one of the reasons why tongue gestures similar to high back vowels are used during sustained note production on the trombone, offsetting the advantages of a minimally constrained open tube that might result from the use of a central vowel tongue position for some of the NZE-speaking participants.

10.4.3 'Laryngeal states' during brass playing

The airflow considerations mentioned above would seem to suggest that brass players should keep their glottis relatively open during playing; however, this is not what empirical studies have found (cf. section 2.3.2 above). The constricted glottal aperture observed in various studies (Mukai, 1992; Dejonckere et al., 1983; King, Ashby, &

Nelson, 1989; Rydell et al., 1996) might nevertheless assist with the regulation of airflow during brass playing by serving as the first step of a two-step mechanism that reduces air pressure on its way from the lungs (the second step being the lip valve emitting the airflow into the mouthpiece; Rydell et al., 1996, pp. 46-47; cf. Yoshikawa, 1998). The question remains, however, whether this mechanism represents active control or whether changes are “self-adjusting or involuntary,” as stated by Bailey (1989, p. 105). That the larynx is capable of subtle interactions with other articulators was shown in a perturbation study by Munhall, Löfqvist and Kelso (1994), although compensations for perturbations at the lips were not always effective (p. 3614). I personally believe that some of the disagreement among the studies listed above might result from the fact that the laryngeal mechanism is comprised of as many as six separate valves (Edmondson & Esling, 2006, p. 159) which might display somewhat independent patterns; I have thus titled this section in agreement with Esling and Harris’ terminological suggestion (2005, p. 378). Furthermore, restricting glottal opening during brass playing (here referring to the true vocal folds) could be a matter of impedance control (5), rather than refer to a fixed glottal distance, allowing for the oscillatory (or co-articulatory) behavior observed by Dejonckere et al. (1983) and Wolfe and Smith (2008; cf. section 2.3.2 above).

10.5 Acoustical considerations related to vocal tract influence on brass instrument sound

All of the above considerations are probably inextricably linked with the resulting vocal tract configurations and, thus, the acoustic impedance of the vocal tract resulting from these constraints. The auditory/acoustic goals of brass playing might, however, provide an additional framework whereby players would try to actively control their vocal tract oscillations, whether consciously or subconsciously, to shape the timbre of the produced sound (cf. section 2.3 above). I agree with Wolfe et al. (2015) that restricting the opening of the true vocal folds (or controlling their impedance - see previous section) not only allows for “fine control of mouth pressure” but also affects potential vocal tract influence considerations by providing a “higher reflection coefficient for acoustic waves in the vocal tract” (p. 3). The result would be a reduced influence of subglottal resonances (cf. section 2.3.2) on the upper vocal tract resonances (extending from the glottis to the lips) interacting with oscillations within the instrument, making it easier for the player to adjust the vocal tract impedance peak falling within the frequency range most strongly emitted from the trombone bell; these considerations support an account of vocal fold constriction during brass playing as deliberate control (cf. previous section). Figure 10.2 (reproduced here for reference from section 2.3.2 – figure 2.11) shows the previously mentioned vocal tract impedance peak around 900 Hz (idealized location for fully closed glottis); in reality, the vocal tract impedance curve during brass playing should pattern somewhere in between the two curves in figure 10.2, due to constricted glottis opening (cf. argumentation in section 2.3.2).

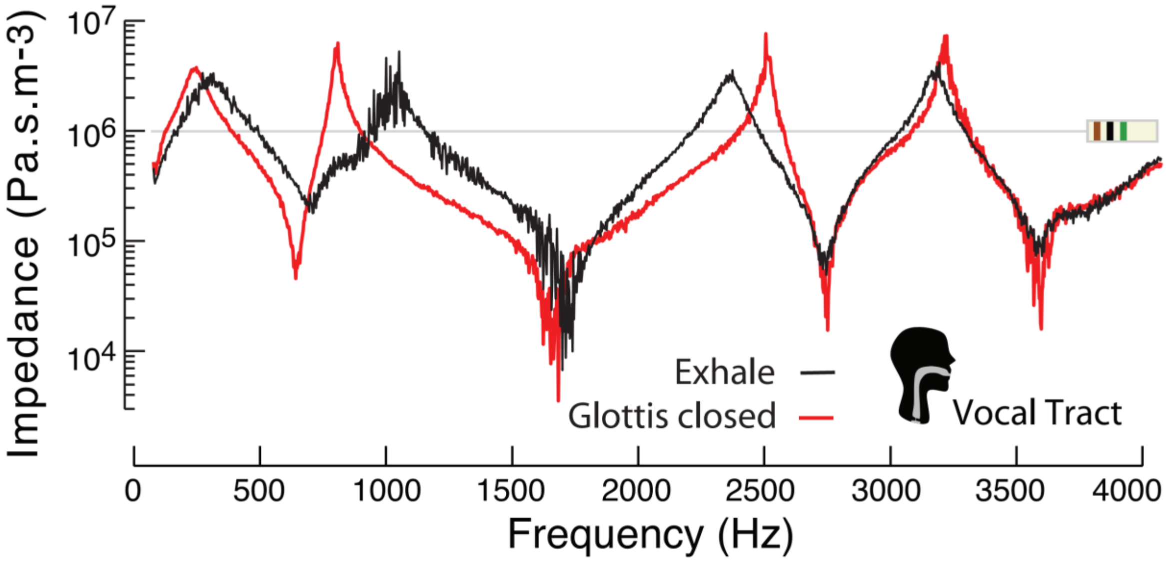


Figure 10.2 (also printed in section 2.3.2): The impedance spectra of a vocal tract measured at the lips: glottis closed (red) and exhaling (black) [original caption]; note that the minima around 650 and 1700 Hz correspond to the formants of [ə] produced by this simple vocal tract configuration (cf. Hanna, Smith, & Wolfe, 2012, p. 3). Reproduced with permission from Wolfe, Fletcher, & Smith, 2015, p. 11; first printed in Hanna, Smith, & Wolfe, 2012.

In section 2.3.2, I concluded that “the combined research findings suggest that the influence of the first vocal tract resonance on brass playing should be limited due to its restricted adjustability”; this first vocal tract resonance is closely linked to glottis opening and thus strongly affected by the considerations presented above. If we now reconsider the frequency range of trombone resonances with its cut-off frequency around 700 Hz (cf. figures 2.3 and 2.5 above), it becomes obvious that, in terms of speech acoustics, a vowel tongue position producing formant frequencies below the cut-off frequency of the trombone (cf. section 2.2) would have the best chance of impacting the timbre of the instrument, as it would have a vocal tract impedance peak at around 4/3 * F1. Higher vocal tract resonances may still affect the timbre at frequencies above the cut-off frequency but their influence would be limited due to the nonlinearities affecting these frequencies (cf. section 2.2 above). In terms of F1, this observation suggests the use of a high tongue position, while for F2 a retracted tongue position would seem advantageous; combining closeness and backing suggests a constraint towards using a high back vowel tongue position, in line with the airflow considerations above. Note however, that it is not yet clear how such a vocal tract maximum affects the sound emitted from the instrument; for the didgeridoo, peaks in the vocal tract impedance spectrum suppress the corresponding harmonics in the radiated sound (Wolfe, Garnier, & Smith, 2009 - Figure 4, p. 14) but the effect is supposedly much weaker for the trombone (cf. section 2.3.1.1 above). I carried out a limited comparison of the sound spectra produced by two players recorded for this thesis for a poster presented at the *International Congress of Phonetic Sciences* in Glasgow (Heyne & Derrick, 2015a); one of the players used a central tongue position (S5 NZE) while the other used a back vowel position (S7 Japanese, not included in this thesis). The findings suggested that a more retracted tongue position leads to “a larger component of high frequencies in the produced sound spectrum” (Heyne & Derrick, 2015a, p. 1); however, this results should not be over-interpreted due to the small sample size and a possible confound in the different horizontal location of the oral constriction produced by the two subjects.

Finally, according to Wolfe et al. (2003), raising the tongue *tip* could also affect vocal tract resonances, by increasing “the height of peaks in the vocal tract impedance,” which would “more effectively” couple vocal tract impedance and instrument resonances (Wolfe et al., 2003, p. 310).

10.6 Motor efficiency considerations affecting the position of the front of the tongue

Several constraints affecting the back of the tongue and the pharyngeal cavity have been discussed above; however, little information has been given regarding the front of the tongue. The limited results for coronal stop consonant production and coronal place of articulation during trombone playing for two participants presented in section 9.5.4 clearly show that the place of articulation during trombone playing is related to coronal stop consonant production in these players' native languages (see figure 9.15). The overall tongue shapes for stop consonants and trombone articulations are also similar for each participant, although the position used by S5 NZE during playing is more different from his speech articulation than for the Tongan participant, especially at the back of the tongue. This could very well be a discrepancy that is influenced by differences in playing proficiency: S4 Tongan is an amateur player while S5 NZE is very skilled semi-professional player. Nonetheless, the coronal place of articulation during trombone playing is clearly influenced by motor memory from the L1 for these two participants; it seems highly likely that the Tongan player produces dental attacks during trombone playing.

I am convinced that the difference regarding place of articulation during trombone playing is also responsible for the group-level difference in tongue height for sustained note productions shown by the z-scored data (cf. figure 9.5). Producing dental attacks seems to require that the tongue be placed more anteriorly during sustained note production than for (presumably) alveolar attacks, reducing the physical distance it has to travel to reach its place of articulation at the beginning (and end) of notes. The much smaller physical distance the NZE player's tongue has to travel could also mean that this player can produce trombone articulations by controlling a single functional segment of the tongue (the front or tip), while the Tongan player would need to recruit several muscle modules to move the whole tongue body, resulting in a much larger physical and, possibly, computational effort.

10.7 Language influence

Secondary to the constraints outlined above, motor memory from one's native language would affect tongue position during trombone playing after a player has selected a locally optimized position based on airflow and acoustical requirements, and considerations related to motor efficiency. If this position shares relevant features with, or spatially approximates, a vowel tongue position for a certain functional segment or the whole shape of the tongue, the motor control system of a player might be inclined to switch to this position or adjust certain aspects of tongue shape to reduce error and muscular economy for subsequent motor movements.

This clearly seems to be the case for the back of the tongue during trombone playing, while the position of the front seems to be indirectly influenced by motor memory from a player's L1 via its coronal place of articulation.

10.8 Further considerations

Not all the points raised above could be sufficiently addressed with the data collected for this thesis; they demonstrate, however, the necessity of considering a large number of variables when accessing the influence of native language on trombone playing. One external finding that I believe can be reasonably claimed to apply to my dataset is the constricted state of the glottis during brass playing, based on the agreement across studies and its implication with several of relevant constraints mentioned above. Before presenting a final discussion focusing on the commonalities among all trombone players included in the study, and the meaningful group differences I believe to be arising due to native language influence, attention should be brought once more to the idiosyncrasies displayed by individual players and their importance in contextualizing the overall results.

The constraints regarding airflow and acoustics tied to vocal tract influence can be satisfied using many different vocal tract configurations (similar to the many-to-one mappings documented for speech production, cf. section 3.3.1), forming a continuum with players at one end who regulate changes in pitch (and potentially loudness as well as timbre) almost exclusively by altering tongue position within the oral cavity (with resulting changes in the lower vocal tract seen as strictly reactionary), and players at the other end who do so without hardly changing oral tongue position, relying instead on the various sphincteric mechanisms of the laryngeal articulator to achieve the required alterations. Other skills might influence such idiosyncratic behavior, as is conceivable in the case of NZE participant S5, who is also a proficient barbershop singer; this subject reported that he could "feel his pharyngeal cavity narrowing when playing high notes" (Heyne & Derrick, 2015b, p. 4) while his ultrasound data show only minimal changes in average tongue position when producing different notes.

10.9 Summary

The above paragraphs have identified the following constraints acting upon tongue position during sustained note production on the trombone:

1) Requirements of airflow favor the use of vocal tract configurations that avoid significant constrictions in the pharyngeal and/or oral cavities; vowel tongue positions satisfying this requirement comprise high back vowels and non-low central vowels (optionally grouped as 'raised' in Esling's (2005) 'laryngeal articulator model').
2) Acoustical considerations suggest that (based on a somewhat constricted glottis determining the location of vocal tract impedance peaks) a retracted (in the classical terminology) tongue position could be advantageous by ensuring that the second vocal tract impedance peak is situated below the cut-off frequency of the trombone and thus can potentially affect the sound spectrum (timbre) released from the instrument.
3) Considerations of motor efficiency favor a small physical distance between the coronal place of articulation used to initiate (and end) notes and the position assumed during sustained note production. Additionally, it might be advantageous if movements between these two positions can be carried out using the muscles of a specific functional section of the tongue only.
4) Language influence via motor memory from a player's native language seems to have a significant impact on the place of articulation used during trombone performance which also affects the position of the front of the tongue during sustained note production.
5) Lastly, motor memory of native language vowel tongue positions seems to represent the least rigorous of the constraints identified in this thesis. Given that a suitable position exists in a player's speech motor repertoire that closely approximates a certain characteristic of the tongue position selected according to the constraints mentioned under the previous points, the motor control system might be inclined to switch to or adapt towards such a position.

All of these constraints place diverging requirements on at least two functionally independent parts of the tongue identified in this thesis, namely the back and front of the tongue. Whether the observed differences in tongue placement among the two groups of trombone players investigated in this thesis lead to perceivable differences in the sound produced by trombone players with different native languages, and what acoustical consequences might result from this difference, however, will have to be investigated in future studies.

10.10 Confounds of this study

10.10.1 National schools of playing and different playing styles

While I cannot claim to have presented conclusive evidence that the observed group differences between Tongan and New Zealand trombone players are only due to the influence of language, I believe that there are good reasons to rule out the influence of national schools of playing and different playing styles as primary factors effecting these differences. Of course, it is likely that the concept of national schools of playing is somewhat interrelated with the notion of language influence on brass playing; it is quite conceivable that, e.g., a (hypothesized) timbral correlate of using a high back vowel tongue position during sustained note production (cf. Heyne & Derrick, 2015a) would provide a beginning player with an auditory cue for arriving at an individually and locally optimized tongue position that results in a similar timbre. Similarly, the use of a particular place of articulation during playing would seem to be recoverable in the acoustic signal from the resulting transient (cf. section 2.6), and players from a specific language group might over time alter their articulatory behavior to enhance such cues. Although somewhat removed from being directly influenced by native language, such cues likely contribute to the different national schools of playing outlined in section 2.4.2 and identified by the respondents of my online questionnaire (see chapter 5).

Most players from both language groups recorded for this study reported that they received instruction in a (British) Brass Band style, and thus should not display any significant differences based on this measure; it should be clear, however, that this does not mean that the input received by any of the participants was in any way controlled for.

When asked informally whether they were ever taught to use any specific speech sound considerations in their brass lessons, a majority of participants demonstrated syllables beginning with their native version of the /t/ sound; players from both groups could be aiming for a target that looks the same (orthographically) in the method books they use (cf. section 2.4.1.2). In reality, however, doing so might perpetuate a playing difference that is already encoded in the player’s native speech motor memory; the intent to follow a common standard thus might actually enhance the influence of native language on playing. Since brass players in most cases do not seem to know what exactly they are doing with their vocal tract musculature (cf. Hiigel, 1967; Irvine, 2003), there may be no way to explicitly teach performing a certain articulatory gesture as it would be mediated by the influence of speech motor memory in every case.

10.10.2 Heterogeneity of the two language groups

While the two language groups observed in this study are quite heterogeneous in terms of playing proficiency, it is unlikely that the reported group differences are solely due to this imbalance, although it could have contributed to the differences in overall tongue height. Generally, more proficient players tend to use a lower tongue position during sustained note production (this is true for five out of nine semi-professional and professional participants observed for this study; two participants in this group used a medium-height tongue position) while less proficient players mostly use relatively high average tongue positions (six out of ten intermediate and amateur participants; one participant in this group used a medium-height tongue position). Futhermore, it is possible that proficiency level may have affected average tongue position during sustained note production in a more indirect way via the motor efficiency considerations linking sustained note position to coronal place of articulation (cf. section 10.6). The more plausible interpretation is that there is a need to restrict the physical distance the tongue has to travel to reach the sustained note position after starting a note, which provides an alternative explanation for the higher position of the front of the tongue employed by the Tongan participants; due to their dental articulations during speech and trombone playing they have to bring their tongue further forward than the NZE participants which leaves less room to lower the tongue rather than just retracting it.

In terms of vocal tract morphology and biomechanics (cf. section 3.3.4), the (presumed, though not measured) underlying heterogeneity of individual vocal tract shapes (across and within both language groups) can be seen to support the interpretation that the influence of native language is at least partially responsible for the group differences reported in the results section. Since players with a variety of individual vocal tract morphologies display similar patterns of articulatory behavior during trombone playing relative to their speech production, speech articulation can be regarded as a sort of normalization process that accounts for such differences.

10.10.3 Ultrasound probe stabilization

The most severe confound of this study concerns the head stabilization used to fix the ultrasound probe underneath the participants' chins. While the jaw brace employed in this study features very good values regarding probe motion and rotaton (cf. section 7.8.2), it nevertheless ties tongue motion to jaw motion which is somewhat problematic when observing two activities for which exact jaw motion patterns are unknown; a review of the existing literature suggests that average jaw opening is similar during speech production and brass playing (cf. section 3.6.1). Yet, a bigger concern was the fact that participants might touch the jaw brace with the instrument tubing while playing the musical passages, potentially compromising the whole of their data; luckily this only happened for a single participant whose data could be corrected using translation and rotation of some of his tongue traces (cf. section 8.7.1). Measures implemented to prevent this from happening included reducing the width of an earlier version of the jaw brace, recording frontal video of participants and reminding them not to touch the jaw brace before and during the trombone playing part of the ultrasound recordings. Note that the requirements of operating a trombone also prevented me from correcting for head movement using a technique such as implemented by Mielke et al. (2005) and Miller and Finch (2011).

10.10.4 Other confounds

A number of additional confounds were mentioned in the text where relevant, a complete list of which is given in table 10.1 on the following two pages.

confound	relevant section	risk	comments
speech production versus trombone playing, including different airflow requirements	whole thesis, 10.4	medium	Since this thesis represents the first attempt at investigating language influence on brass playing using articulatory data, no information is currently available to evaluate this confound. Using an imaging method that would allow the observation of a larger section and/or the full dimensionality of the vocal tract, such as MRI, or measuring tongue/jaw coarticulation when using UTI could reduce this confound; however, there are other challenges associated with such techniques, as laid out in section 2.5. In terms of airflow, it should be clear that the two activities place different demands on vocal tract shape (cf. section 10.4 above) and this in turn likely increases muscle stiffness along the vocal tract; even for speech production, however, the effect of different aerodynamic pressures is not yet fully understood (cf. Mooshammer, Hoole, & Kühnert, 1995; Hoole,1998; Fuchs et al., 2004) so that it is currently impossible to properly evaluate this confound.
Tongan versus NZE participants & language	whole thesis, 10.10.2	low	While there might be morphological differences of the vocal tract (these were not measured) between and within the two language groups observed in this study, speech articulation can be regarded as a sort of normalization process limiting this confound.
national schools/styles of playing & different playing styles	10.10.1	medium	The concept of national schools of playing is likely somewhat interrelated with the notion of language influence on brass playing, and it is possible that some or all of the characteristics of a certain style of playing used by a specific group of musicians are acquired by imitation. The close patterning of average sustained note position with selected aspects of speech articulation documented in this thesis, however, suggests that language influence is the driving factor behind the documented group-level differences. Schools/styles of playing can certainly can aggravate this influence, and e.g., jazz playing might be more susceptible to allow such influence as it places greater importance on an individual ‘sound’ compared to Brass Band and orchestral styles of playing.
difference in trombone playing proficiency levels (between language groups)	10.10.2	medium	Although the two language groups observed in this study are quite heterogeneous in terms of playing proficiency, it is unlikely that the observed group differences are solely due to this imbalance. It is true that more proficient players are more likely to use a lower tongue position during sustained note production, and the group including a higher number of professional and semi-professional players are

			the NZE participants who also feature a lower average tongue position for sustained note productions; there exists an alternative explanation for the overall tongue height difference, however, which arises out of motor efficiency requirements (cf. section 10.6) regarding the place of articulation used during trombone playing.
difference of experimental materials, due to limited proficiency in Tongan	7.4.2	low	My limited expertise in Tongan meant that I had to use a dictionary to compile the wordlist for the Tongan participants, relying on the fact that Tongan spelling is mostly phonetic. The result was the inclusion of a large number of English loanwords, which may have made the articulatory results for the two language groups more similar, if at all affecting the articulations produced by the Tongan participants.
different recording locations	7.6	medium	The room used for recordings in Tonga was bigger than the room used at the University of Canterbury and did not feature sound attenuation. Sound from a nearby workshop, a gym, chickens, and the more reverberant acoustics may have affected the Tongan participants' speech production and trombone playing, and there was no chance to control for room temperature. These confounds, however, are much smaller than failing to control for different mouthpiece and instrument shapes, which were the same for all subjects except the pilot participant; the confound is further mitigated by the fact that I did not carry any acoustic measurements on the data recorded in Tonga or used it for perceptual judgment tasks.
ultrasound probe stabilization	7.8.2	high	The jaw brace employed in this study ties tongue motion to jaw motion which is somewhat problematic when observing activities for which exact jaw motion patterns are unknown; the existing literature, however, suggests that average jaw opening is similar during speech production and brass playing (cf. section 3.6.1). More seriously, there was the danger that participants might touch the jaw brace with the instrument tubing, potentially compromising the recorded data. A number of measures were implemented to prevent this from happening, such as reducing the width of the jaw brace, recording frontal video of participants and reminding them not to touch the jaw brace before and during the trombone playing part of the ultrasound recordings.

Table 10.1: List of confounds for this study.

10.11 Implications

The findings of this study show that two activities that have previously been linked through their cognitive mechanisms, language and music, can also be connected more indirectly via motor memory resulting from a shared physiological system. Although both activities are clearly forms of communication, one is inherently non-referential (if we disregard vocal music with lyrics), while the other is by definition referential or semiotic (but see Bowling et al., 2010; Curtis & Bharucha, 2010 for findings challenging this traditional distinction).

My findings also support modular accounts of motor control and provide relevant suggestions for the pedagogy of brass instruments, as outlined in the following sections.

10.11.1 Implications for modular theories of motor control

The finding that different sections of the tongue are subject to, and react differently to constraints arising from various sources provides support for modular accounts of motor control (cf. section 3.2.3), especially as proposed for the facial-oral-laryngeal-respiratory musculature (cf. section 3.3.3). The fact that most predictions listed in section 3.7 were borne out by the results of my research suggests that muscle synergies do not need to be encoded phylogenetically or pertain to “innate, spontaneous actions” (Mayer et al., in press, para. 2; cf. the term motor ‘primitives’) in order to become retrievable for other complex motor behaviors (cf. the concept of ‘negative transfer’ in Second Language Acquisition research; e.g., Colantoni, Steele, & Escudero, 2015). The transfer of muscle modules – alternatively conceivable as motor memory – from speech production to brass playing rather constitutes transfer of muscle synergies from one complex acquired form of motor behavior to another. It is hoped that my findings will help to stir further research (e.g., on other wind instruments, whistling, or the mouth harp) to fill the void of literature on the neuromuscular modularization of speech production and other functions of the upper vocal tract (Gick & Stavness, 2013, p. 1).

10.11.2 Implications for brass pedagogy

The ubiquity of the use of speech syllables in brass teaching (cf. section 2.4), along with the notion of national schools of playing (cf. section 10.10 above), provided a major impetus to empirically investigate the influence of native language on brass playing. Prior to this thesis, a number of studies were able to show that most brass players do not actually use the vowel tongue positions associated with the speech syllables recommended in many brass method books (cf. Irvine, 2003). One of my initial assumptions was that maybe prior studies had not looked at the right kind of vowel tongue shapes (e.g., schwa /ə/; cf. section 2.4.1.3) to make the connection between them and the tongue shapes used during brass playing, or that they had not collected enough data to find such an effect. While a few individual players in this study use a schwa-like position during sustained note production, this assumption was not supported on the group level. Does the finding that the tongue shape during sustained note production on the trombone patterns differently at the back and front thus discredit the use of speech syllables in brass pedagogy?

If we think of the process of finding a tongue shape to use during sustained note production on the trombone as local optimization or coming up with a 'good-enough' solution, then maybe it does not matter that these tongue shapes differ to a certain extent. It could be that the higher-order constraints specified above overwrite many of the specifications that come with a pre-defined vowel tongue position, meaning that only a subset of the muscle activations, e.g., associated with the vowel /i/ in the syllable /ti/, are actually realized by a brass player when given the instruction to replicate this vocal tract shape during playing. Effectively, the use of speech syllables might thus serve to get brass players to explore various local maxima in the relevant articulatory space and abandon inferior motor habits associated with a previous local optimization process. Given the limited proprioceptive innervation of the tongue and the lack of alternatives to specify tongue position, this might actually be the most efficient way of reaching such a goal (cf. Loeb, 2012, p. 764). Nevertheless, I would encourage brass teachers to use syllables that more accurately reflect the tongue position players are likely to employ.

The reader should also be aware that while I believe the constraints listed above to apply to all brass instruments, further research is needed to determine how the constraints arising from airflow and acoustical requirements, and motor efficiency, might affect the playing of brass instruments with different bore lengths (e.g., trumpet versus tuba), taper (e.g., trombone versus euphonium), and mouthpiece size/shape in relation to bore length (e.g., French horn). These considerations in turn could lead to an increased or decreased role of native language influence on sound production on these instruments when compared to the trombone players investigated in this study.

10.12 Future directions

Limited results of articulation data for two participants presented in section 9.5.4 promise a fruitful line of research, given the large differences associated with motor efficiency requirements in covering the distance between the place of articulation and the position assumed during sustained notes. Note, however, that the current result could be somewhat overstated due to differences in participants' playing proficiencies. Recently, a new version of the GetContours script for MATLAB has become available (Tiede & Whalen, 2015), which implements automatic contour tracking; this could render the required tracing of multiple ultrasound frames for each articulatory gesture less tedious.

Another idea concerns measuring the articulatory distance between the tongue shape for each note and the relevant average curve for that pitch, using the area measurement technique explicated in section 9.4, and comparing these results with measurements of the timbre for each individual note and the average timbre for a certain pitch. Carrying out such measurements for various sections of the tongue should provide insights regarding the influence of tongue height and/or fronting on the timbre produced by trombone players.

Brass playing also offers a unique opportunity to investigate articulator kinematics when subject to aerodynamic pressures exceeding those generated during speech production. An appropriate visualization technique to use for such research would be electromagnetic articulography (section 2.5.5). While limited research has investigated such effects during speech production (Mooshammer, Hoole, & Kühnert, 1995; Hoole, 1998; Fuchs et al., 2004), I propose using a continuum that incorporates speech at different intensities (cf. Munhall, Ostry, & Flanagan, 1991; Hoole 1998), brass playing, and possibly singing and other wind instruments requiring lesser air pressures, to advance the understanding of articulator kinematics.

11 Conclusion

This thesis attempted to answer the question whether there is an influence of native language on the playing of brass instruments. Two hypotheses were formulated on the basis of a comprehensive review of previous empirical research on brass playing (the biggest to date), and studies of motor control and speech production. Hypothesis 1 about the perceptibility of playing differences among players with different native languages was answered cautiously in the affirmative by an online questionnaire completed by 135 respondents world-wide; it was acknowledged, however, that further research is needed to properly address this question. Hypothesis 2 predicted that the tongue position assumed during sustained note production would be based on motor memory from a player's native language vowel production. To address this hypothesis, ten Tongan and nine New Zealand English-speaking trombone players were recorded using midsagittal ultrasound of the tongue while reading wordlists and playing the trombone. While no match was found between the overall tongue shape for vowels and sustained note production on the language group level (although this was true for a small number of individual players), different patterns were found to govern the behavior at the back and front of the tongue. No support was thus provided for hypothesis 2a) that predicted sustained note tongue shape would resemble vowel tongue shape; however, an alternative option (hypothesis 2b)) predicted that functionally independent sections of the tongue would be individually affected by motor memory from a player's native language. Language influence was suggested to be responsible for a patterning of sustained note shape with the back vowels in each language at the back of the tongue, while motor efficiency considerations seemed to contribute to a large positional difference at the front of the tongue. Overall, language influence was found to be secondary to more basic constraints arising from airflow and acoustical requirements, and considerations of articulatory efficiency. A number of confounds were listed in the discussion section, the most serious of which concern the difficult task of fixing an ultrasound sound place underneath a trombone player's chin during trombone playing, the challenge of comparing articulatory movements during two very different activities, and differences in trombone playing proficiency levels of the participants comprising the two observed language groups.

The findings of this study provide strong support for modular theories of motor control, and upper vocal tract control, specifically, by showing that motor memory from an acquired, highly-skilled behavior (speech production) can influence another skilled behavior (trombone playing).

The procedure used to normalize midsagittal ultrasound data across individuals and different activities outlined in the analysis section furthermore constitutes an important contribution to ultrasound methodology.

References

Adachi, S., & Sato, M. (1996). Trumpet sound simulation using a two-dimensional lip vibration model. *The Journal of the Acoustical Society of America, 99*, 1200.

Adatia, A., & Gehring, E. (1971). Proprioceptive innervation of the tongue. *Journal of Anatomy, 110*(Pt 2), 215.

Agricola, M. (1896). *Musica instrumentalis deudsch, 1. und 4. Ausgabe Wittenberg 1528 und 1545. In neuer diplomatisch genauer, zum Teil faksimilierter Ausgabe*. Leipzig: Breitkopf & Härtel.

Aldred, H. M. (1997). *Ifi palasa: Brass bands in Tonga* (Unpublished Master of Music dissertation). Victoria University of Wellington, Wellington, New Zealand.

Alfonso, P., & Van Lieshout, P. (1997). Spatial and temporal variability in obstruent gestural specification by stutterers and controls: Comparisons across sessions. In W. Hulstijn, H. F. M. Peters, & P. Van Lieshout (Eds.), *Speech production: Motor control, brain research and fluency disorders* (pp. 151–160). Amsterdam, the Netherlands: Elsevier.

Alibiglou, L., & MacKinnon, C. D. (2012). The early release of planned movement by acoustic startle can be delayed by transcranial magnetic stimulation over the motor cortex. *The Journal of Physiology, 590*(4), 919-936.

Alimurung, G. (2014, July 3) Regrets? This 100-year-old man has a few. Retrieved from http://www.laweekly.com/arts/regrets-this-100-year-old-man-has-a-few-4824124

Altenburg, J. E. (1974). *Essay on an introduction to the heroic and musical trumpeters's and kettledrummers' art* (E. Tarr, Trans. & Ed.). Nashville, TN: The Brass Press. (Original work published 1795)

Altenburg, J. E. (1993). *Versuch einer Anleitung zur heroisch-musikalischen Trompeter-und Pauker-Kunst*. Hofheim & Leipzig: Friedrich Hofmeister. (Original work published 1795)

Amazeen, P. G. (2002). Is dynamics the content of a generalized motor program for rhythmic interlimb coordination? *Journal of Motor Behavior, 34*(3), 233-251.

Amerman, J. D., Daniloff, R., & Moll, K. L. (1970). Lip and jaw coarticulation for the phoneme /æ/. *Journal of Speech, Language, and Hearing Research, 13*(1), 147-161.

Amstutz, A. K. (1970). *A videofluorographic study of the teeth aperture instrument pivot and tongue arch and their influence on trumpet performance* (Unpublished doctoral dissertation), University of Oklahoma, Norman, OK.

Amstutz, K. A. (1977). A videofluorographic study of the teeth aperture, instrument pivot and tongue arch and their influence on trumpet performance. *Journal of the International Trumpet Guild, 2*, 25-26.

Anderson, V., & Otsuka, Y. (2006). The Phonetics and Phonology of" Definitive Accent" in Tongan. *Oceanic Linguistics*, 21-42.

Antão, G., & Moreira, D. (2015). *Intercultural musical learning in the era of technology*. Paper presented at the 3rd 21st Century Academic Forum Conference at Harvard, Boston, MA.

Arban, J.-B. (1936). *Complete conservatory method for trumpet (cornet) or Eb alto, Bb bass in treble clef*. New York: Carl Fischer.

Arbeau, T. (1967). *Orchésographie* (M. Stewart Evans, Trans.). New York: Dover. (Original work published 1588)

Arbeau, T. (1972). *Orchésographie*. Geneva: Minkoff. (Original work published 1588)

Articulate Instruments, Ltd. (2008). Ultrasound Stabilisation Headset Users Manual: Revision 1.4. Edinburgh, UK: Articulate Instruments Ltd.

Ayers, R. D., Eliason, L. J., & Mahgerefteh, D. (1985). The conical bore in musical acoustics. *American Journal of Physics, 53*(6), 528-537.

Baayen, R., Piepenbrock, R., & Gulikers, L. (1995). CELEX2 LDC96L14 [database]. Philadelphia: Linguistic Data Consortium.

Backhaus, W. (1932). Die Ausgleichsvorgänge in der Akustik. *Zeitschrift für technische Physik, 13*, 31-46.

Backus, J. (1983). Effect of the player's vocal tract on the tone quality of reed woodwind instruments. *The Journal of the Acoustical Society of America, 74*(S1), S29.

Baer, T. (1979). Reflex activation of laryngeal muscles by sudden induced subglottal pressure changes. *The Journal of the Acoustical Society of America, 65*(5), 1271-1275.

Baer, T., Gore, J. C., Gracco, L., & Nye, P. (1991). Analysis of vocal tract shape and dimensions using magnetic resonance imaging: Vowels. *The Journal of the Acoustical Society of America, 90*(2), 799-828.

Bailey, R. E. (1989). *An investigation of the laryngeal activity of trumpet players during the performance of selected exercises* (Unpublished doctoral dissertation). University of North Texas, Denton, TX.

Barboteu, G. (2000). The evolution of the horn in france and its school (Nancy Fako, Trans.) *The Horn Call - Journal of the International Horn Society, 30*(4), 33-38. (Original work published in 1975)

Barenboim, D., & Said, E. W. (2004). *Parallels and paradoxes: Explorations in music and society*. New York: Random House.

Barras, C., Geoffrois, E., Wu, Z., & Liberman, M. (2001). Transcriber: Development and use of a tool for assisting speech corpora production. *Speech Communication, 33*(1), 5-22.

Barry, W., & Andreeva, B. (2001). Cross-language similarities and differences in spontaneous speech patterns. *Journal of the International Phonetic Association, 31*(1), 51-66.

Bauer, L., & Warren, P. (2004). New Zealand English: Phonology. In E. W. Schneider, K. Burridge, B. Kortmann, R. Mesthrie, & C. Upton (Eds.), *A Handbook of Varieties of English* (Vol. 1, pp. 580-602). Berlin: Mouton de Gruyter.

Beauchamp, J. W. (2012). Trombone Transfer Functions: Comparison Between Frequency-Swept Sine Wave and Human Performer Input. *Archives of Acoustics, 37*(4), 447-454.

Beddor, P. S. (2015). The relation between language users' perception and production repertoires. In The Scottish Consortium for ICPhS 2015 (Eds.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, Scotland.

Benade, A. (1969). Effect of dispersion and scattering on the startup of brass instrument tones. *The Journal of the Acoustical Society of America, 45*(1), 296-297.

Benade, A. H. (1976). *Fundamentals of Musical Acoustics*. London: Oxford University Press.

Benade, A. H. (1978). The physics of brasses. In C. M. Hutchins (Ed.), *The Physics of Music, Readings from Scientific American* (pp. 44-55). San Francisco: WH Freeman & Co.

Benade, A. H. (1985). Air column, reed, and player's windway interaction in musical instruments. In I. R. Titze & R. C. Scherer (Eds.), *Vocal Fold Physiology, Biomechanics, Acoustics and Phonatory Control* (pp. 425-452): Denver Center for the Performing Arts, Denver, CO.

Benade, A. H. (1986). Interactions between the player’s windway and the air column of a musical instrument. *Cleveland Clinic Quarterly, 53*, 27-32.

Benade, A., Greenberg, P., Hoekje, P., & Suttle, R. (1977). Wind instrument very-nearfield and room-average spectra. *The Journal of the Acoustical Society of America, 61*(S1), S35.

Bendinelli, C. (1975). *The entire art of trumpet playing* (E. Tarr Ed.). Nashville, TN: The Brass Press. (Original manuscript 1614)

Benedetti, F. (1988). Localization of tactile stimuli and body parts in space: two dissociated perceptual experiences revealed by a lack of constancy in the presence of position sense and motor activity. *Journal of Experimental Psychology: Human Perception and Performance, 14*(1), 69.

Bernstein, N. A. (1967). *The co-ordination and regulation of movements*. Oxford, UK: Pergamon Press.

Berio, L. (2006). Sequenza V: for trombone solo (1966). Vienna: Universal Edition.

Bertsch, M. A. (2013). *Tonguing on brass instruments: Tempo and endurance*. Paper presented at the International Symposium on Performance Science (ISPS), Vienna, Austria.

Bertsch, M., & Hoole, P. (2014). *Tonguing on brass instruments - Highspeed visualization and benchmarks of fastest tempi*. Paper presented at the International Symposium on Musical Acoustics (ISMA), Le Mans, France.

Bianco, T., Fréour, V., Cossette, I., Bevilacqua, F., & Causse, R. (2012). Measures of Facial Muscle Activation, Intra-oral Pressure and Mouthpiece Force in Trumpet Playing. *Journal of New Music Research, 41*(1), 49-65.

Bianco, T., Fréour, V., Rasamimanana, N., Bevilaqua, F., & Caussé, R. (2010). On gestural variation and coarticulation effects in sound control. In S. Kopp & I. Wachsmuth (Eds.), *Gesture in embodied communication and human-computer interaction* (pp. 134-145). Heidelberg: Springer.

Bismantova, B. (1978). The Discussion of Wind Instruments in B. Bismantova's Compendio musicale (1677): Translation and Commentary (B. Dickey, P. Leonards, & E. Tarr, Trans.). In B. Dickey, P. Leonards, & E. Tarr (Eds.), *Basler Jahrbuch für historische Musikpraxis* (Vol. II). Winterthur, Switzerland: Amadeus.

Bizzi, E., & Cheung, V. C. (2013). The neural origin of muscle synergies. *Frontiers in computational neuroscience, 7*, 51.

Boersma, P., & Weenink, D. (2014). Praat: Doing phonetics by computer [Computer program]. Version 5.3.52. Retrieved May 18, 2014.

Bouchard, K. E., Mesgarani, N., Johnson, K., & Chang, E. F. (2013). Functional organization of human sensorimotor cortex for speech articulation. *Nature, 495*(7441), 327-332.

Boutin, H., Fletcher, N., Smith, J., & Wolfe, J. (2013). *Lip motion, the playing frequency of the trombone and the upstream and downstream impedances*. Paper presented at the Stockholm Music Acoustics Conference (SMAC), Stockholm, Sweden.

Boutin, H., Fletcher, N., Smith, J., & Wolfe, J. (2015). Relationships between pressure, flow, lip motion, and upstream and downstream impedances for the trombone. *The Journal of the Acoustical Society of America, 137*(3), 1195-1209.

Bowling, D. L., Gill, K., Choi, J. D., Prinz, J., & Purves, D. (2010). Major and minor music compared to excited and subdued speech. *The Journal of the Acoustical Society of America, 127*(1), 491-503.

Bowman, J. (2011). Trumpet Pedagogy Challenges in Thailand. *Journal of the International Trumpet Guild*, 70, 90.

Bressmann, T., Thind, P., Uy, C., Bollig, C., Gilbert, R. W., & Irish, J. C. (2005). Quantitative three-dimensional ultrasound analysis of tongue protrusion, grooving and symmetry: Data from 12 normal speakers and a partial glossectomee. *Clinical Linguistics & Phonetics, 19*(6-7), 573-588.

Bromage, S., Richards, O., & Campbell, D. (2003). *Reproducibility and control of the embouchure of an artificial mouth for playing brass instruments.* Paper presented at the Stockholm Musical Acoustics Conference (SMAC), Stockholm.

Bromage, S., Campbell, M., & Gilbert, J. (2010). Open areas of vibrating lips in trombone playing. *Acta Acustica united with Acustica, 96*(4), 603-613.

Browman, C. P., & Goldstein, L. M. (1986). Towards an articulatory phonology. *Phonology, 3*(01), 219-252.

Brunner, J., Fuchs, S., & Perrier, P. (2009). On the relationship between palate shape and articulatory behavior. *The Journal of the Acoustical Society of America, 125*(6), 3936-3949.

Buchaillard, S., Perrier, P., & Payan, Y. (2009). A biomechanical model of cardinal vowel production: Muscle activations and the impact of gravity on tongue positioning. *The Journal of the Acoustical Society of America, 126*(4), 2033-2051.

Budde, P. J. (2011). *An analysis of methods for teaching middle school band students to articulate* (Unpublished doctoral dissertation). University of Minnesota, Minneapolis, MN.

Campbell, M., & Greated, C. (1994). *The musician's guide to acoustics*. Oxford: Oxford University Press.

Campbell, M., Myers, A., & Chick, J. (2013). Influence of the bell profile of the trombone on sound reflection and radiation. *Proceedings of Meetings on Acoustics, 19*(1).

Cardano, G. (1973). *Writings on music* (C. A. Miller, Trans.): American Institute of Musicology. (Original work published 1663)

Carleton, A. (1938). Observations on the problem of the proprioceptive innervation of the tongue. *Journal of Anatomy, 72*, 502-507.

Carlsen, A. N., Maslovat, D., & Franks, I. M. (2012). Preparation for voluntary movement in healthy and clinical populations: Evidence from startle. *Clinical Neurophysiology, 123*(1), 21-33.

Carral, S. (2011). Determining the just noticeable difference in timbre through spectral morphing: A trombone example. *Acta Acustica united with Acustica, 97*(3), 466-476.

Carral, S., & Campbell, M. (2002). *The influence of the mouthpiece throat diameter on the perception of timbre of brass instruments.* Paper presented at the International Symposium on Musical Acoustics (ISMA), Mexico City, Mexico.

Carter, W. A. (1969). The Role of the Glottis in Brass Playing. *Brass Anthology: A compendium of articles from the Instrumentalist on playing the brass instruments.*, 425-428. (Original work published 1966)

Caussé, R., Kergomard, J., & Lurton, X. (1984). Input impedance of brass musical instruments—comparison between experiment and numerical models. *The Journal of the Acoustical Society of America, 75*(1), 241-254.

Chen, J.-M., Smith, J., & Wolfe, J. (2008). How to play the first bar of Rhapsody in Blue. *Journal of the Acoustical Society of America, 123*(5), 3123.

Chen, J.-M., Smith, J., & Wolfe, J. (2011). Saxophonists tune vocal tract resonances in advanced performance techniques. *The Journal of the Acoustical Society of America, 129*(1), 415-426.

Chen, J.-M., Smith, J., & Wolfe, J. (2012). Do trumpet players tune resonances of the vocal tract? *The Journal of the Acoustical Society of America, 131*(1), 722-727.

Cheng, H. Y., Murdoch, B. E., Goozee, J. V., & Scott, D. (2007). Electropalatographic assessment of tongue-to-palate contact patterns and variability in children, adolescents, and adults. *Journal of Speech, Language, and Hearing Research, 50*(2), 375-392.

Chesky, K., Devroop, K., & Ford, J. (2002). Medical problems of brass instrumentalists: Prevalence rates for trumpet, trombone, French horn, and low brass. *Medical Problems of Performing Artists, 17*(2), 93-98.

Chi, X., & Sonderegger, M. (2007). Subglottal coupling and its influence on vowel formants. *The Journal of the Acoustical Society of America, 122*, 1735.

Chick, J., Logie, S., Campbell, M., & Gilbert, J. (2012). *Spectral enrichment and wall losses in trombones played at high dynamic levels.* Paper presented at the Acoustics 2012.

Chick, J., Logie, S., Norman, L., & Murray, C. (2013). Transient phenomena in brass instruments. *Proceedings of Meetings on Acoustics, 19*.

Chiu, C., & Gick, B. (2014). Pitch planning in English and Taiwanese Mandarin: evidence from startle-elicited responses. *The Journal of the Acoustical Society of America, 136*(4), EL322-EL328.

Chiu, C., Stevenson, A., Maslovat, D., Chua, R., Gick, B., & Franks, I. (2011). Feed-forward control of phonetic gestures in consonant–vowel syllables: Evidence from responses to auditory startle. *The Journal of the Acoustical Society of America, 129*(4), 2455.

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics, 27*(2), 207-229.

Christensen, R. H. B. (2015). Package 'ordinal': Regression Models for Ordinal Data [Computer program]. Version 2015.6-28.

Churchward, C. M. (1953). *Tongan grammar*. Oxford: Oxford University Press.

Chvatal, S. A., & Ting, L. H. (2013). Common muscle synergies for balance and walking. *Frontiers in computational neuroscience, 7*, 48.

Cilingir, H. Z. (2012). *The Relationship of oral anatomy and trumpet performance: Prediction of physical talent* (Unpublished doctoral dissertation), University of Cincinnati.

Clinch, P., Troup, G., & Harris, L. (1982). The importance of vocal tract resonance in clarinet and saxophone performance, a preliminary account. *Acta Acustica united with Acustica, 50*(4), 280-284.

Clivio, G. P. (1996). Italian. In T. J. McGee, A. G. Rigg, & D. N. Klausner (Eds.), *Singing Early Music: The pronunciation of European languages in the late Middle Ages and Renaissance* (pp. 187-211). Bloomington, IN: Indiana University Press.

Coelho, F. M., & Iazzetta, F. (2011). *Measuring the Influence of the Vocal Tract in the Timbre of the Flute: Preliminary Aspects*. Paper presented at the 18th International Congress on Sound and Vibration, Rio de Janeiro, Brazil.

Colantoni, L., Steele, J., & Escudero, P. (2015). *Second Language Speech*. Cambridge: Cambridge University Press.

Coltman, J. W. (1973). Mouth resonance effects in the flute. *The Journal of the Acoustical Society of America, 54*(2), 417-420.

Copeman, H. (1996). Italian Latin. In T. J. McGee, A. G. Rigg, & D. N. Klausner (Eds.), *Singing Early Music: The pronunciation of European languages in the late Middle Ages and Renaissance* (pp. 212-215). Bloomington, IN: Indiana University Press.

Copley, D. C., & Strong, W. J. (1996). A stroboscopic study of lip vibrations in a trombone. *The Journal of the Acoustical Society of America, 99*(2), 1219.

Cox, F., & Palethorpe, S. (2008). *Reversal of short front vowel raising in Australian English.* Paper presented at the 9th Annual Conference of the International Speech Communication Association, Brisbane, Australia.

Cox, K. A. (2014). *Connections between linguistic and musical sound systems of British and American trombonists* (Unpublished doctoral dissertation). The University of North Carolina at Greensboro, Greensboro, NC.

Crees, E., & Gane, P. (1988). *How trombonists do it: A book on trombone legato: Lip flexibilities – Warm ups – Studies (Bass Clef edition).* Manton, England: Brass Wind Publications.

Crystal, T. H., & House, A. S. (1988). A note on the variability of timing control. *Journal of Speech, Language, and Hearing Research, 31*(3), 497-502.

Cullen, J., Gilbert, J., & Campbell, D. (2000). Brass instruments: Linear stability analysis and experiments with an artificial mouth. *Acta Acustica united with Acustica, 86*(4), 704-724.

Curtis, M. E., & Bharucha, J. J. (2010). The minor third communicates sadness in speech, mirroring its use in music. *Emotion, 10*(3), 335.

Dalla Casa, G. (1970). *Il vero modo di diminuir: Con tutte le sorti di stromenti [di fiato, & corda, & di voce humana]* (G. Vecchi Ed.). Bologna: Forni. (Original work published 1584)

Dang, J., & Honda, K. (2002). Estimation of vocal tract shapes from speech sounds with a physiological articulatory model. *Journal of Phonetics, 30*(3), 511-532.

d'Avella, A., Giese, M., Ivanenko, Y. P., Schack, T., & Flash, T. (2015). Editorial: Modularity in motor control: From muscle synergies to cognitive action representation. *Frontiers in computational neuroscience, 9*.

Davidson, L. (2006). Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *The Journal of the Acoustical Society of America, 120*(1), 407-415.

Dawson, K. M., Tiede, M. K., & Whalen, D. (2015). Methods for quantifying tongue shape and complexity using ultrasound imaging. *Clinical Linguistics & Phonetics*, 1-17.

Dejonckere, P., Orval, F., Miller, R., & Sneppe, R. (1983). Mécanisme oscillatoire de la glotte dans le jeux de cor. *Brass Bulletin, 41*, 28-35.

Derrick, D., Best, C. T., & Fiasson, R. (2015). Non-metallic ultrasound probe holder for co-collection and co-registration with EMA. In The Scottish Consortium for ICPhS 2015 (Eds.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, Scotland.

Devroop, K., & Chesky, K. (2002). *Comparison of biomechanical forces generated during trumpet performance in contrasting settings*. Paper presented at the 20th Annual Symposium on the Medical Problems of Musicans and Dancers, Aspen, CO.

Deye, H. W. (1947). The Use of the Tongue in Brass Instruments. *Brass Anthology*(March/April), 11.

De Young, D. D. (1975). *A videofluorographic analysis of the pharyngeal opening during performance of selected exercises for trombone* (Unpublished doctoral dissertation). University of Minnesota., Minneapolis, MN.

Diaz, F. (2011). The effect of attack consonants on perception of intensity in trombone onsets. *Journal of Band Research, 46*(2).

Diehl, R. L., Lindblom, B., Hoemeke, K. A., & Fahey, R. P. (1996). On explaining certain male-female differences in the phonetic realization of vowel categories. *Journal of Phonetics, 24*(2), 187-208.

Dominici, N., Ivanenko, Y. P., Cappellini, G., d'Avella, A., Mondì, V., Cicchese, M., ... Giannini, C. (2011). Locomotor primitives in newborn babies and their development. *Science, 334*, 997-999.

Dovel, J. (2009). The Diphthong versus the pure vowel in trumpet performance and pedagogy. *Journal of the International Trumpet Guild, 29*, 54, 62.

Edmondson, J. A., & Esling, J. H. (2006). The valves of the throat and their functioning in tone, vocal register and stress: laryngoscopic case studies. *Phonology, 23*(2), 157-191.

Elliott, C. A. (1975). Attacks and releases as factors in instrument identification. *Journal of Research in Music Education, 23*(1), 35-40.

Elliott, S. J., & Bowsher, J. M. (1982). Regeneration in brass wind instruments. *Journal of Sound and Vibration, 83*(2), 181-217.

Epstein, E. K. [MRI Horn Videos: Pedagogy Informed by Science] (2016, June 24). Eli Epstein: Finger Breathing. [Video file]. Retrieved from https://www.youtube.com/watch?v=Vjhe-JkFPrk

Epstein, E. K., & Iltis, P. [MRI Horn Videos: Pedagogy Informed by Science] (2016a, June 24). Episode 1: Introduction to the RT-MRI Horn Project. [Video file]. Retrieved from https://www.youtube.com/watch?v=LTMe38uOaT8&list=PL5mAKZsdj5P3OS6c4LMyTJmW9WCUs31UL

Epstein, E. K., & Iltis, P. [MRI Horn Videos: Pedagogy Informed by Science] (2016b, June 24). Episode 2: The Role of the Tongue and Jaw in Pitch Placement. [Video file]. Retrieved from https://www.youtube.com/watch?v=7Cz5HoQ1fCI&index=3&list=PL5mAKZsdj5P3OS6c4LMyTJmW9WCUs31UL

Epstein, M. A., & Stone, M. (2005). The tongue stops here: Ultrasound imaging of the palate. *The Journal of the Acoustical Society of America, 118*(4), 2128-2131.

Epstein, M. A., Stone, M., Pouplier, M., & Parthasarathy, V. (2004). Obtaining a palatal trace for ultrasound images. *The Journal of the Acoustical Society of America, 115*(5), 2631-2632.

Erickson, R. (1969). General speech: For trombone solo (with theatrical effects). Baltimore, MD: Smith publications.

Esling, J. H. (2005). There are no back vowels: The laryngeal articulator model. *The Canadian journal of linguistics/La revue canadienne de linguistique, 50*(1), 13-44.

Esling, J. H., & Harris, J. G. (2005). States of the glottis: An articulatory phonetic model based on laryngoscopic observations. In W. Hardcastle & J. Beck (Eds.), *A figure of speech: A Festschrift for John Laver* (pp. 347-383). Mahwah, NJ: Lawrence Erlbaum Associates.

Esling, J. H., Moisik, S. R., & Crevier-Buchman, L. (2016). *Putting the larynx in the vowel space: Studying larynx state across vowel quality using MRI*. Paper presented at the Colloquium of the British Association of Academic Phoneticians, Lancaster, England.

Fang, Q., Fujita, S., Lu, X., & Dang, J. (2009). A model-based investigation of activations of the tongue muscles in vowel production. *Acoustical Science and Technology, 30*(4), 277-287.

Fant, G. (1960). *Acoustic theory of speech production with calculations based on X-ray studies of Russian articulations* (Vol. 2). The Hague: Walter de Gruyter.

Fantini, G. (1978). *Method for learning to play the trumpet in a warlike way as well as musically: with the organ, with a mute, with a harpsichord, and every other instrument...* (E. Tarr, Trans.). Nashville, TN: The Brass Press. (Original work published 1638)

Farkas, P. (1956). *The art of French horn playing*. New York: Alfred Music.

Farnetani, E., & Faber, A. (1992). Tongue-jaw coordination in vowel production: Isolated words versus connected speech. *Speech Communication, 11*(4-5), 401-410.

Ferragne, E., & Pellegrino, F. (2010). Formant frequencies of vowels in 13 accents of the British Isles. *Journal of the International Phonetic Association, 40*(1), 1-34.

FFmpeg. (2014). FFmpg [computer program]. https://ffmpeg.org/ Retrieved November 19, 2014.

Fisher, P. K., Schmidlin, S. L., Douglass, N., & Schultheis, J. (2014). *Oral musculation patterns influencing speech, voice, and instrument tone*. Paper presented at the International Association of Oral Myologists, Calgary, Canada.

Fitch, W. T., & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *The Journal of the Acoustical Society of America, 106*(3), 1511-1522.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology, 47*(6), 381-391.

Fitzgerald, B. (1946). Articulation. *Brass Anthology* (November/December), 3.

Fletcher, N. H. (1983). Acoustics of the Australian didjeridu. *Australian Aboriginal Studies*(1), 28-37.

Fletcher, N. H., & Rossing, T. D. (1991). *The Physics of Musical Instruments.* New York: Springer.

Foulkes, P., & Docherty, G. (2014). *Urban voices: Accent studies in the British Isles*. London: Routledge.

Fowler, C. A., Rubin, P., Remez, R. E., & Turvey, M. (1980). Implications for speech production of a general theory of action. In B. Butterworth (Ed.), *Language production* (pp. 373-420). New York: Academic Press.

Fox, F. (1982). *Essentials of brass playing (1974)*. Pittsburgh: Volkwein Bros.

Frederiksen, B. (2006). *Arnold Jacobs: Song and wind* (J. Taylor, Ed.). Gurnee, IL: Windsong Press.

Frenzel, P. (1996). Late Medieval German and Early New High German. In T. J. McGee, A. G. Rigg, & D. N. Klausner (Eds.), *Singing Early Music: The pronunciation of European languages in the late Middle Ages and Renaissance* (pp. 243-257). Bloomington, IN: Indiana University Press.

Fréour, V. (2013). *Acoustic and respiratory pressure control in brass instrument performance* (Unpublished doctoral dissertation), McGill University, Montreal, Canada.

Fréour, V., Lopes, N., Hélie, T., Causse, R., & Scavone, G. (2015). In-vitro and numerical investigations of the influence of a vocal-tract resonance on lip auto-oscillations in trombone performance. *Acta Acustica united with Acustica, 101*(2), 256-269.

Fréour, V., & Scavone, G. (2010). *Vocal-tract influence in trombone performance*. Paper presented at the International Symposium on Music Acoustics (ISMA), Sydney & Katoomba, Australia.

Fréour, V., & Scavone, G. P. (2013). Acoustical interaction between vibrating lips, downstream air column, and upstream airways in trombone performance. *The Journal of the Acoustical Society of America, 134*(5), 3887-3898.

Frère, J., & Hug, F. (2012). Between-subject variability of muscle synergies during a complex motor skill. *Frontiers in Computational Neuroscience, 6*, 99.

Friston, K. (2011). What is optimal about motor control? *Neuron, 72*(3), 488-498.

Fritz, C., Caussé, R., Kergomard, J., & Wolfe, J. (2005). *Experimental study of the influence of the clarinettist's vocal tract.* Paper presented at the Proc. Forum Acusticum.

Fritz, C., & Wolfe, J. (2005). How do clarinet players adjust the resonances of their vocal tracts for different playing effects? *The Journal of the Acoustical Society of America, 118*(5), 3306-3315.

Fritz, C., Wolfe, J., Kergomard, J., & Caussé, R. (2003). *Playing frequency shift due to the interaction between the vocal tract of the musician and the clarinet.* Paper presented at the Stockholm Music Acoustics Conference (SMAC).

Frohrip, K. R. (1972). *A videofluorographic analysis of certain physiological factors involved in performance of selected exercises for trombone* (Unpublished doctoral dissertation), University of Minnesota., Minneapolis, MN.

Fromont, R., & Hay, J. (2012). *LaBB-CAT: An annotation store.* Paper presented at the Australasian Language Technology Association Workshop.

Frucht, S. J. (2009). Embouchure dystonia—Portrait of a task-specific cranial dystonia. *Movement disorders, 24*(12), 1752-1762.

Fuchs, S., Hoole, P., Brunner, J., & Inoue, M. (2004). *The trough effect: An aerodynamic phenomenon.* Paper presented at the CD-ROM MIT-meeting: From Sound to Sense.

Fuchs, S., Perrier, P., Geng, C., & Mooshammer, C. (2006). What role does the palate play in speech motor control? Insights from tongue kinematics for German alveolar obstruents. In J. Harrington & M. Tabain (Eds.), *Speech production: Models, phonetic processes, and techniques* (pp. 149-164). New York: Psychology Press.

Fuchs, S., Winkler, R., & Perrier, P. (2008). *Do speakers' vocal tract geometries shape their articulatory vowel space?* Paper presented at the 8th International Seminar on Speech Production (ISSP).

Ganassi, S. (1956). *Opera intitulata Fontegara: Venice 1535: a treatise on the art of playing the recorder and of free ornamentation* (H. Peter, Trans.). Berlin: Robert Lienau. (Original work published 1535)

Ganassi, S. (1969). *Opera intitulata Fontegara.* Bologna: Forni. (Original work published 1535)

Ganesh, G., Haruno, M., Kawato, M., & Burdet, E. (2010). Motor memory and local minimization of error and effort, not global optimization, determine motor behavior. *Journal of Neurophysiology, 104*(1), 382-390.

Garellek, M., & White, J. (2015). Phonetics of Tongan stress. *Journal of the International Phonetic Association, 45*(1), 13-34.

Gérard, J.-M., Wilhelms-Tricarico, R., Perrier, P., & Payan, Y. (2006). A 3D dynamical biomechanical tongue model to study speech motor control (Original paper published 2003). *arXiv preprint physics/0606148*.

Gibson, D. J. (1942). *Recordings of differentiations in tone color produced by vowel sounds on the trumpet* (Unpublished MA dissertation). University of Minnesota., Minneapolis, MN.

Gick, B. (2002). An X-ray investigation of pharyngeal constriction in American English schwa. *Phonetica, 59*(1), 38-48.

Gick, B. (2016). Ecologizing Dimensionality: Prospects for a Modular Theory of Speech Production. *Ecological psychology, 28*(3), 176-181.

Gick, B., Allen, B., Roewer-Despres, F., & Stavness, I. (in press). Speaking tongues are actively braced. *Journal of Speech, Language & Hearing Research*.

Gick, B., Anderson, P., Chen, H., Chiu, C., Kwon, H. B., Stavness, I., ... Fels, S. (2014). Speech function of the oropharyngeal isthmus: a modelling study. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 2*(4), 217-222.

Gick, B., & Derrick, D. (2009). Aero-tactile integration in speech perception. *Nature, 462*(7272), 502-504.

Gick, B., & Stavness, I. (2013). Modularizing speech. *Frontiers in Psychology, 4*, 977.

Gick, B., Stavness, I., Chiu, C., & Fels, S. (2011). Categorical variation in lip posture is determined by quantal biomechanical-articulatory relations. *Canadian Acoustics, 39*(3), 178-179.

Gick, B., Wilson, I., & Derrick, D. (2013). *Articulatory phonetics*. Malden, MA: Wiley-Blackwell.

Gick, B., Wilson, I., Koch, K., & Cook, C. (2005). Language-specific articulatory settings: Evidence from inter-utterance rest position. *Phonetica, 61*(4), 220-233.

Gilbert, J., Ruiz, L. L., & Gougeon, S. (2006). *Influence de la température sur la justesse d'un instrumenta vent.* Paper presented at the Proceedings of Congres Français d'Acoustique, Tours, France.

Gordon, E., Campbell, L., Hay, J., Maclagan, M., Sudbury, A., & Trudgill, P. (2004). *New Zealand English: Its origins and evolution*. Cambridge: Cambridge University Press.

Gordon, E., Maclagan, M., & Hay, J. (2007). The ONZE corpus. In J. C. Beal, K. P. Corrigan, & H. L. Moisl (Eds.), *Creating and digitizing language corpora* (Vol. 2: Diachronic databases, pp. 82-104). New York: Springer.

Grover, C., & Craske, B. (1991). The effect of loading on position sense in the tongue. *Perception & Psychophysics, 50*(1), 7-14.

Gu, C. (2013a). Package 'gss': A comprehensive package for structural multivariate function estimation uwsing smoothing splines. [Computer program]. Version 2.1-0.

Gu, C. (2013b). Smoothing spline ANOVA models (Vol. 297). New York: Springer Science & Business Media.

Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review, 102*(3), 594.

Hageman, J. E. (2005). *The Horn: A discussion of the nationalistic schools of horn performance and the players and composers who influenced them* (Unpublished doctoral dissertation). University of California at Los Angeles, Los Angeles, CA.

Hall, J. C. (1954). *A radiographic, spectrographic, and photographic study of the non-labial physical changes which occur in the transition from middle to low and middle to high registers during trumpet performance* (Unpublished doctoral dissertation). Indiana University, Bloomington, IN.

Hanna, N., Smith, J., & Wolfe, J. (2012). *Low frequency response of the vocal tract: acoustic and mechanical resonances and their losses.* Paper presented at the Annual Conference of the Australian Acoustical Society, Fremantle, Australia.

Hanna, N., Smith, J., & Wolfe, J. (2016). Frequencies, bandwidths and magnitudes of vocal tract and surrounding tissue resonances, measured through the lips during phonation. *The Journal of the Acoustical Society of America, 139*(5), 2924-2936.

Hanneton, S., Berthoz, A., Droulez, J., & Slotine, J.-J. E. (1997). Does the brain use sliding variables for the control of movements? *Biological cybernetics, 77*(6), 381-393.

Hanson, H. M., & Stevens, K. N. (1995). *Sub-glottal resonances in female speakers and their effect on vowel spectra.* Paper presented at the 13th International Congress of Phonetic Sciences, Stockholm, Sweden.

Harshman, R., Ladefoged, P., & Goldstein, L. (1977). Factor analysis of tongue shapes. *The Journal of the Acoustical Society of America, 62*(3), 693-707.

Härtel, K. (2012). In aller Munde: Die Zunge. *Clarino, February*, 20-23.

Hartwell, L. H., Hopfield, J. J., Leibler, S., & Murray, A. W. (1999). From molecular to modular cell biology. *Nature, 402*, C47-C52.

Haslinger, B., Altenmüller, E., Castrop, F., Zimmer, C., & Dresel, C. (2010). Sensorimotor overactivity as a pathophysiologic trait of embouchure dystonia. *Neurology, 74*(22), 1790-1797.

Hauf, R. (2007). In aller Munde: Gedanken zur Funktionsweise der Zunge bei Blechbläsern. *clarino.print praxis, April*.

Hay, J., Maclagan, M., & Gordon, E. (2008). *New Zealand English*. Edinburgh: Edinburgh University Press.

Haynie, J. J. (1969). *A Videofluorographic Presentation of the Physiological Phenomena Influencing Trumpet Performance.* North Texas State University: School of Music.

Haynie, J. J., & Hardin, A. F. (2007). *Inside John Haynie's studio: A master teacher's lessons on trumpet and life*. Denton, TX: University of North Texas Press.

Hélie, T., Lopes, N., & Caussé, R. (2012). Open-loop control of a robotized artificial mouth for brass instruments. *The Journal of the Acoustical Society of America, 131*(4), 3470-3470.

Hélie, T., Vergez, C., Lévine, J., & Rodet, X. (1999). *Inverse problem in a physical model of trumpet.* Paper presented at the International Computer Music Conference (ICMC).

Helmholtz, H. L. (1954). *On the Sensations of Tone as a Physiological Basis for the Theory of Music (1877)* (A. J. Ellis, Trans. 2nd ed.). New York: Dover Publications.

Henrich, N., d'Alessandro, C., Doval, B., & Castellengo, M. (2005). Glottal open quotient in singing: Measurements and correlation with laryngeal mechanisms, vocal intensity, and fundamental frequency. *The Journal of the Acoustical Society of America, 117*(3), 1417-1430.

Henry, F. M. (1968). Specificity vs. generality in learning motor skill. In R. C. Brown & G. S. Kenyon (Eds.), *Classical studies on physical activity* (pp. 331-340). Englewood Cliffs, NJ: Prentice Hall. (Original publication 1958)

Hering, S. (2011). For Progressive Etudes for Trombone (Revised Edition). New York: Carl Fischer.

Heuer, H. (1991). Invariant relative timing in motor-program theory. In J. Fagard & P. H. Wolff (Eds.), *The development of timing control and temporal organization in coordinated action* (pp. 37-68). Amsterdam: Elsevier.

Heyne, M. (2015). *The 'singing approach' to brass playing revisited: We sing the language we speak*. Paper presented at the 2nd Conference of the Australian Music & Psychology Society, Western Sydney University, Australia.

Heyne, M., & Derrick, D. (2014). *Some initial findings regarding first language influence on playing brass instruments.* Paper presented at the 15th Australasian International Conference on Speech Science and Technology (SST), Christchurch, New Zealand.

Heyne, M., & Derrick, D. (2015a). The influence of tongue position on trombone sound: A likely area of language influence. In The Scottish Consortium for ICPhS 2015 (Eds.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, Scotland.

Heyne, M., & Derrick, D. (2015b). *Trombone players seem to use different tongue positions while playing sustained notes, depending on their native languages*. Poster presented at the 9th Triennial Conference of the European Society for the Cognitive Sciences of Music (ESCOM), Manchester, England.

Heyne, M., & Derrick, D. (2015c). Using a radial ultrasound probe's virtual origin to compute midsagittal smoothing splines in polar coordinates. *Journal of the Acoustical Society of America, 138*(6), EL509-514.

Heyne, M., & Derrick, D. (2015d). *Why expert listeners might be able to distinguish differences in sound produced by brass players from different language backgrounds.* Poster presented at the 10th UK Language Variation and Change Conference (UKLVC), York, England.

Heyne, M., & Derrick, D. (2016a). *An articulatory and acoustic study of monophthongs in New Zealand English*. Paper presented at the New Zealand English and English in New Zealand Conference (NZEENZ), University of Canterbury, Christchurch, New Zealand.

Heyne, M., & Derrick, D. (2016b). Visualization techniques for empirical brass instrument research. *Journal of the International Trumpet Guild, 40*, 6-14, 24.

Heyne, M., & Derrick, D. (2016c). Stylistic variation of sub-phonemic syllabic influences on New Zealand English schwa production. Paper to be presented at New Ways of Analyzing Variation 45, Simon Fraser University, Vancouver, Canada.

Hézard, T., Fréour, V., Caussé, R., Hélie, T., & Scavone, G. P. (2014). Synchronous multimodal measurements on lips and glottis: Comparison between two human-valve oscillating systems. *Acta Acustica united with Acustica, 100*(6), 1172-1185.

Hiigel, L. E. (1967). *The relationship of syllables to pitch and tonguing in brass instrument playing* (Unpublished doctoral dissertation). University of California at Los Angeles, Los Angeles, CA.

Hiki, S., & Itoh, H. (1986). Influence of palate shape on lingual articulation. *Speech Communication, 5*(2), 141-158.

Hirata, Y., Schulz, M., Altenmüller, E., Elbert, T., & Pantev, C. (2004). Sensory mapping of lip representation in brass musicians with embouchure dystonia. *Neuroreport, 15*(5), 815-818.

Hirayama, M., Vatikiotis-Bateson, E., Kawato, M., & Jordan, M. I. (1992). Forward dynamics modeling of speech motor control using physiological data. In R. P. Lippman, J. E. Moody, & D. S. Touretsky (Eds.), *Advances in neural information processing systems, Vol. 4* (pp. 191-198). San Mateo, CA: Morgan Kauffmann Publishers.

Hirschberg, A., Gilbert, J., Msallam, R., & Wijnands, A. (1996). Shock waves in trombones. *The Journal of the Acoustical Society of America, 99*(3), 1754-1758.

Honda, K., Maeda, S., Hashi, M., Dembowski, J., & Westbury, J. R. (1996). *Human palate and related structures: Their articulatory consequences.* Paper presented at the 4th International Conference on Spoken Language Processing, Philadelphia, PA.

Honda, K., Takano, S., & Takemoto, H. (2010). Effects of side cavities and tongue stabilization: Possible extensions of the quantal theory. *Journal of Phonetics, 38*(1), 33-43.

Honikman, B. (1964). Articulatory settings. In D. Abercrombie, D. B. Fry, P. A. D. MacCarthy, N. C. Scott, & J. L. M. Trim (Eds.), *In honour of Daniel Jones: Papers contributed on the occasion of his eightieth birthday, 12 September 1961* (pp. 73-84). London: Longman.

Hoole, P. (1998). Do airstream mechanisms influence tongue movement paths? *Phonetica, 55*(3), 131-146.

Hoole, P., & Munhall, K. (1994). Do air-stream mechanisms influence tongue movement paths? *The Journal of the Acoustical Society of America, 95*, 2821.

Houde, R. (1968). A study of tongue body motion during selected consonant sounds. *Speech Communications Research Laboratory, Santa Barbara, SCRL Monograph, 2*.

Hruška, V., Švejda, M., & Guštar, M. (2015). *Transients of the trumpet tone: Basic links between perception and measurements of lips opening area and pressure in player's mouth.* Paper presented at the Third Vienna Talk on Music Acoustics, Vienna.

Iltis, P. W., Frahm, J., Voit, D., Joseph, A., Schoonderwaldt, E., & Altenmüller, E. (2015a). Divergent oral cavity motor strategies between healthy elite and dystonic horn players. *Journal of Clinical Movement Disorders, 2*(1), 1.

Iltis, P. W., Frahm, J., Voit, D., Joseph, A. A., Schoonderwaldt, E., & Altenmüller, E. (2015b). High-speed real-time magnetic resonance imaging of fast tongue movements in elite horn players. *Quantitative Imaging in Medicine and Surgery, 5*(3), 374-381.

Iltis, P. W., Schoonderwaldt, E., Zhang, S., Frahm, J., & Altenmüller, E. (2015c). Real-time MRI comparisons of brass players: A methodological pilot study. *Human Movement Science, 42*, 132-145.

Iltis, P. W., Frahm, J., Voit, D., & Joseph, A. (2016). Inefficiencies in Motor Strategies of Horn Players with Embouchure Dystonia. *Medical Problems of Performing Artists, 31*(2), 69-77.

Imagawa, H., Kiritani, S., Masaki, S., & Shirai, K. (1985). Contextual variation in the jaw position for the vowels in /CVC/ utterances. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics, 19*, 7-19.

Irvine, G. (2003). Spotlight on brass: Lessons learned (and ignored) from the past. *Canadian Winds: The Journal of the Canadian Band Association, 2*(1), 17-19.

Johnson, K. (2011). *Acoustic and auditory phonetics*. Malden, MA: John Wiley & Sons.

Johnson, K., Ladefoged, P., & Lindau, M. (1993). Individual differences in vowel production. *The Journal of the Acoustical Society of America, 94*, 701.

Johnston, R., Clinch, P., & Troup, G. (1986). The role of the vocal tract resonance in clarinet playing. *Acoustics Australia, 14*(3), 67-69.

Jun, S. A., Lee, S. h., Beckman, M. E., & De Jong, K. (1991). The role of the jaw in consonant articulation. *The Journal of the Acoustical Society of America, 90*(4), 2362-2363.

Jung, Y. (2009). *Acoustic articulatory evidence for quantal vowel categories: the features [low] and [back]* (Unpublished doctoral dissertation), Massachusetts Institute of Technology, Cambridge, MA.

Kaburagi, T., Yamada, N., Fukui, T., & Minamiya, E. (2011). A methodological and preliminary study on the acoustic effect of a trumpet player's vocal tract. *The Journal of the Acoustical Society of America, 130*(1), 536-545.

Kakita, Y., Fujimura, O., & Honda, K. (1985). Computation of mapping from muscular contraction patterns to formant patterns in vowel space. In V. A. Fromkin (Ed.), *Phonetic Linguistics: Essays in honor of Peter Ladefoged* (pp. 157-168). Orlando: Academic Press.

Keating, P. A., Lindblom, B., Lubker, J., & Kreiman, J. (1994). Variability in jaw height for segments in English and Swedish VCVs. *Journal of Phonetics, 22*(4), 407-422.

Keating, P. A., Mikoś, M. J., & Ganong III, W. F. (1981). A cross-language study of range of voice onset time in the perception of initial stop voicing. *The Journal of the Acoustical Society of America, 70*(5), 1261-1271.

Keele, S. W., & Hawkins, H. L. (1982). Explorations of individual differences relevant to high level skill. *Journal of Motor Behavior, 14*(1), 3-23.

Keele, S. W., Ivry, R. I., & Pokorny, R. A. (1987). Force control and its relation to timing. *Journal of Motor Behavior, 19*(1), 96-114.

Keele, S. W., Pokorny, R. A., Corcos, D. M., & Ivry, R. (1985). Do perception and motor production share common timing mechanisms: A correlational analysis. *Acta psychologica, 60*(2-3), 173-191.

Kelso, J. A., Saltzman, E. L., & Tuller, B. (1986). The dynamical perspective on speech production: Data and theory. *Journal of Phonetics, 14*(1), 29-59.

Kendall, R., & Carterette, E. (1989). *Perceptual, verbal, and acoustical attributes of wind instrument dyads.* Paper presented at the First International Conference on Music Perception and Cognition (ICMPC), Kyoto, Japan.

Kendall, R. A., & Carterette, E. C. (1993a). Verbal attributes of simultaneous wind instrument timbres: I. von Bismarck's adjectives. *Music Perception: An Interdisciplinary Journal, 10*(4), 445-468.

Kendall, R. A., & Carterette, E. C. (1993b). Verbal Attributes of Simultaneous Wind Instrument Timbres: II. Adjectives Induced from Piston's" Orchestration". *Music Perception: An Interdisciplinary Journal, 10*(4), 469-502.

Kendall, R., & Carterette, E. (1996). *Difference thresholds for timbre related to spectral centroid.* Paper presented at the 4th International Conference on Music Perception and Cognition (ICMPC), Montreal, Canada.

Kennedy, G. D., & Deverson, T. (2005). *The New Zealand Oxford Dictionary*. Auckland: Oxford University Press.

Kern, M., Birn, R., Jaradeh, S., Jesmanowicz, A., Cox, R., Hyde, J., & Shaker, R. (2001). Swallow-related cerebral cortical activity maps are not specific to deglutition. *American Journal of Physiology-Gastrointestinal and Liver Physiology, 280*(4), G531-G538.

Kiesling, S. F. (2005). Variation, stance and style: Word-final -er, high rising tone, and ethnicity in Australian English. *English World-Wide, 26*(1), 1-42.

King, A. I., Ashby, J., & Nelson, C. (1989). Laryngeal function in wind instruments: The brass. *Journal of Voice, 3*(1), 65-67.

Kiritani, S., Tanaka, T., Hashimoto, K., Masaki, S., & Shirai, K. (1983). Contextual variation of the jaw movement for the intervocalic consonant in VCV utterances. *Ann. Bull. RILP, 17*, 45-53.

Klapp, S. T. (1996). Reaction time analysis of central motor control. In H. Zelaznik (Ed.), *Advances in motor learning and control* (pp. 13-35). Champaign, IL: Human Kinetics.

Klapp, S. T. (2003). Reaction time analysis of two types of motor preparation for speech articulation: Action as a sequence of chunks. *Journal of Motor Behavior, 35*(2), 135-150.

Klatt, D. H. (1975). Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech, Language, and Hearing Research, 18*(4), 686-706.

Krivokapić, J. (2012). Prosodic planning in speech production. In S. Fuchs, M. Weihrich, D. Pape, & P. Perrier (Eds.), *Speech planning and dynamics. Peter Lang* (pp. 157-190): Peter Lang.

Kronman, U., & Sundberg, J. (1987). Is the musical ritard an allusion to physical motion. *Action and perception in rhythm and music, 55*, 57-68.

Kudo, K., & Ohtsuki, T. (2008). Adaptive variability in skilled human movements. *Information and Media Technologies, 3*(2), 409-420.

Kühnert, B., Ledl, C., Hoole, P., & Tillmann, H. (1991). Tongue-jaw interactions in lingual consonants. *PERILUS (Phonetic Experimental Research at the Institute of Linguistics University of Stockholm), 14*, 21-25.

Kula, K., Cilingir, H. Z., Eckert, G., Dagg, J., & Ghoneima, A. (2015). The association of malocclusion and trumpet performance. *The Angle orthodontist, 86*(1), 108-114.

Kuo, G., & Vicenik, C. J. (2012). The Intonation of Tongan. *UCLA Working Papers in Phonetics, 111*, 63-91.

Kuttruff, H. (2014). *Room acoustics* (5th revised ed.). Abingdon: CRC Press.

Laboissière, R., Schwartz, J.-L., & Bailly, G. (1991). Modelling the speaker-listener interaction in a quantitative model for speech motor control: a framework and some preliminary results, PERILUS XIV. *PERILUS (Phonetic Experimental Research at the Institute of Linguistics University of Stockholm), 14*, 57-62.

Lacquaniti, F., Ivanenko, Y. P., D'avella, A., Zelik, K., & Zago, M. (2013). Evolutionary and developmental modules. *Frontiers in Computational Neuroscience, 7*, 61.

Ladd, D. R., Turnbull, R., Browne, C., Caldwell-Harris, C., Ganushchak, L., Swoboda, K., ... Dediu, D. (2013). Patterns of individual differences in the perception of missing-fundamental tones. *Journal of Experimental Psychology: Human Perception and Performance, 39*(5), 1386.

LaGorio, L. A., Carnaby-Mann, G. D., & Crary, M. A. (2008). Cross-system effects of dysphagia treatment on dysphonia: a case report. *Cases journal, 1*(1), 1.

Laguitton, V., Demany, L., Semal, C., & Liégeois-Chauvel, C. (1998). Pitch perception: a difference between right-and left-handed listeners. *Neuropsychologia, 36*(3), 201-207.

Lammert, A., Proctor, M., & Narayanan, S. (2013a). Interspeaker variability in hard palate morphology and vowel production. *Journal of Speech, Language, and Hearing Research*, 56(6), S1924-S1933.

Lammert, A., Proctor, M., & Narayanan, S. (2013b). Morphological variation in the adult hard palate and posterior pharyngeal wall. *Journal of Speech, Language, and Hearing Research*, 56(2), 521-530.

Lane, H., & Webster, J. W. (1991). Speech deterioration in postlingually deafened adults. *The Journal of the Acoustical Society of America, 89*(2), 859-866.

Laquemant, J., & Chrestian, J. (1998). *Recueil de pièces de viole en musique et en tablature* (S. Cheney Ed.). Geneva: Editions Minkoff. (Original work published 1666)

Larson, A. (n. d.). Trombone lesson: Brass embouchure with vowels. Retrieved from http://www.digitaltrombone.com/trombone-embouchure-using-vowels.html\

Laver, J. (1978). The concept of articulatory settings: a historical survey. *Historiographia Linguistica, 5*(1-2), 1-14.

Lawson, E., & Mills, L. (2014). *Using ultrasound tongue imaging to study variation in the GOOSE vowel across accents of English*. Paper presented at Methods in Dialectology XV, Groningen, Netherlands.

Lawson, E., Mills, L., & Stuart-Smith, J. (2015). *Variation in tongue and lip movement in the GOOSE vowel across British Isles Englishes*. Paper presented at the 10th UK Language Variation and Change Conference (UKLVC), York, England.

Lebrecht, Norman. (2016, June 30). ‘I was victim of an oboe hate mob’. Retrieved from http://slippedisc.com/2016/06/i-was-victim-of-an-oboe-hate-mob/

Leno, H. L. (1970). *Lip vibration characteristics of selected trombone performers* (Unpublished doctoral dissertation). University of Arizona, Tucson, AZ.

Leung, Wai Kit. (n. d.) [Review of the CD Dues for oboe and piano, by K. Needleman & J. Lim] Retrieved from http://musicweb-international.com/classrev/2016/Apr/Oboe_duos_GEN16407.htm

Lewis, M. P. (Ed.). (2009). *Ethnologue: Languages of the world* (Vol. 16). Dallas: SIL international.

Li, W., Chen, J.-M., Smith, J., & Wolfe, J. (2015). Effect of vocal tract resonances on the sound spectrum of the saxophone. *Acta Acustica united with Acustica, 101*(2), 270-278.

Liberman, A. M., & Whalen, D. H. (2000). On the relation of speech to language. *Trends in cognitive sciences, 4*(5), 187-196.

Lindblad, P., & Lundqvist, S. (1999). *How and why do the tongue gestures of [t],[d],[l],[n],[s], and [r] differ?* Paper presented at the 14th International Congress of Phonetic Sciences (ICPhS), San Francisco, CA.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20*(3), 384-422.

Lisker, L., & Abramson, A. S. (1967). Some effects of context on voice onset time in English stops. *Language and Speech, 10*(1), 1-28.

Loeb, G. E. (2012). Optimal isn’t good enough. *Biological Cybernetics, 106*(11-12), 757-765.

Löfqvist, A., & Gracco, V. L. (2002). Control of oral closure in lingual stop consonant production. *The Journal of the Acoustical Society of America, 111*(6), 2811-2827.

Löfqvist, A., & Yoshioka, H. (1984). Intrasegmental timing: Laryngeal-oral coordination in voiceless consonant production. *Speech Communication, 3*(4), 279-289.

Logie, S., Chick, J., Stevenson, S., & Campbell, M. (2010). *Upward and downward slurred transients on brass instruments: Why is one not simply the inverse of the other?* Paper presented at the 10ème Congrès Français d'Acoustique.

Loubriel, L. (2011a). *Brass Singers: The teaching of Arnold Jacobs*. Chicago: Scholar Publications.

Loubriel, L. (2011b). *Lasting change for trumpeters* (2nd printing): Minneapolis, MN, & Chicago, IL: Scholar Publications.

Ludwigsen, D. O. (2003). *Physical modeling of the trombone player's lips*. Paper presented at the Stockholm Music Acoustics Conference (SMAC), Stockholm, Sweden.

Lulich, S. M. (2010). Subglottal resonances and distinctive features. *Journal of Phonetics, 38*(1), 20-32.

Lulich, S. M. (2014). Combined analysis of real-time three-dimensional tongue ultrasound and digitized three-dimensional palate impressions: Methods and findings. *The Journal of the Acoustical Society of America, 136*(4), 2104-2104.

Lulich, S. M. (2016). Patterns of lingual articulation: A real-time three-dimensional ultrasound+ palate study. *The Journal of the Acoustical Society of America, 139*(4), 2192-2192.

Lulich, S. M., Alwan, A., Arsikere, H., Morton, J. R., & Sommers, M. S. (2011). Resonances and wave propagation velocity in the subglottal airways. *The Journal of the Acoustical Society of America, 130*(4), 2108-2115.

Lulich, S. M., Charles, S., & Lulich, B. (2016). Real-time three-dimensional tongue motion during clarinet performance. *The Journal of the Acoustical Society of America, 139*(4), 2203-2203.

Maclagan, M., & Hay, J. (2004). The rise and rise of New Zealand English DRESS. In S. Cassidy, F. Cox, R. Mannell, & S. Palethorpe (Eds.), *Proceedings of the 10th Australian speech science and technology conference* (pp. 183-188). Sydney: Australian Speech Science and Technology Association Inc.

MacNeilage, P. F. (2008). *The origin of speech*. Oxford: Oxford University Press.

MacNeilage, P. F., & Sholes, G. N. (1964). An electromyographic study of the tongue during vowel production. *Journal of Speech, Language, and Hearing Research, 7*(3), 209-232.

Malas, K., Trudeau, N., Chagnon, M., & McFarland, D. H. (2015). Feeding–swallowing difficulties in children later diagnosed with language impairment. *Developmental Medicine & Child Neurology, 57*(9), 872-879.

Martin, D. W. (1942). Lip vibrations in a cornet mouthpiece. *The Journal of the Acoustical Society of America, 13*(3), 305-308.

Martin, R. E., MacIntosh, B. J., Smith, R. C., Barr, A. M., Stevens, T. K., Gati, J. S., & Menon, R. S. (2004). Cerebral areas processing swallowing and tongue movement are overlapping but distinct: A functional magnetic resonance imaging study. *Journal of Neurophysiology, 92*(4), 2428-2493.

MathWorks, Inc. (2015). MATLAB 2015a [computer program]. Natick, MA.

Matthews, P. (1964). Muscle spindles and their motor control. *Physiological Reviews, 44*(2), 219-288.

Mayer, C., Roewer-Despres, F., Stavness, I., & Gick, B. (in press). Do innate stereotypies serve as a basis for swallowing and learned speech movements? *Behavioral and Brain Sciences.*

McAuliffe, M. J., Ward, E. C., & Murdoch, B. E. (2001). Tongue-to-palate contact patterns and variability of four English consonants in an /i/ vowel environment. *Logopedics Phoniatrics Vocology, 26*(4), 165-178.

McCann, J. L. (1989). *A history of trumpet and cornet pedagogy in the United States, 1840-1942* (Unpublished doctoral dissertation), Northwestern University, Evanston, IL.

McFarland, D. H., & Tremblay, P. (2006). Clinical implications of cross-system interactions. *Seminars in speech and language, 27*(4), 300-309.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*, 746-748.

Meidt, J. A. (1967). *A cinefluorographic investigation of oral adjustments for various aspects of brass instrument performance* (Unpublished doctoral dissertation), The University of Iowa, Iowa City, IA.

Ménard, L., Aubin, J., Thibeault, M., & Richard, G. (2011). Measuring tongue shapes and positions with ultrasound imaging: A validation experiment using an articulatory model. *Folia Phoniatrica et Logopaedica, 64*(2), 64-72.

Merriman, L. C., & Meidt, J. A. (1968). A cinefluorographic investigation of brass instrument performance. *Journal of Research in Music Education, 16*(1), 31-38.

Mersenne, M. (1957). *Harmonie universelle V: The Book on Instruments* (R. Chapman, Trans.). The Hague: M. Nijhoff. (Original work published 1636)

Mersenne, M. (1965). *Harmonie universelle V*. Paris: Editions du Centre National de la Recherche Scientifique. (Original work published 1636)

Mesthrie, R. (2008). Synopsis: The phonology of english in africa and south and southeast asia. In R. Mesthrie (Ed.), *Varieties of English: Africa, South and Southeast Asia* (pp. 307-322). New York: Mouton de Gruyter.

Mielke, J. (2013). tongue_ssanova.r: Functions for SSANOVA comparisons of tongue traces in polar coordinates using gss *retrieved from http://phon.chass.ncsu.edu/manual/tongue_ssanova.r*.

Mielke, J. (2015). An ultrasound study of Canadian French rhotic vowels with polar smoothing spline comparisonsa). *The Journal of the Acoustical Society of America, 137*(5), 2858-2869.

Mielke, J., Baker, A., Archangeli, D., & Racy, S. (2005). Palatron: A technique for aligning ultrasound images of the tongue and palate. In D. Siddiqi & B. V. Tucker (Eds.), (Vol. 14, pp. 97-108).

Miller, A. L., & Finch, K. B. (2011). Corrected high–frame rate anchored ultrasound with software alignment. *Journal of Speech, Language, and Hearing Research, 54*(2), 471-486.

Mills, L. (2014). *An articulatory study of goose-fronting in varieties of English* (Unpublished MSc dissertation), Queen Margaret University, Edinburgh, UK.

Milner, P. [YouTube Symphony Orchestra 2011] (2010, October 20). LSO Master Class – Bass Trombone. [Video file]. Retrieved from https://www.youtube.com/watch?v=-JKF526GGzc

Miyawaki, K., Hirose, H., Ushijima, T., & Sawashima, M. (1975). A preliminary report on the electromyographic study of the activity of lingual muscles. *Ann. Bull. RILP, 9*(91), 406.

Moerman, K. M. (2014). interp_polar [MATLAB function, 23 Sept 2014]. https://github.com/Kevin-Mattheus-Moerman/GIBBON/blob/master/lib/interp_polar.m.

Moisik, S., & Gick, B. (2013). The quantal larynx revisited. *Proceedings of Meetings on Acoustics, 19*.

Moore, B. C. (2003). *An introduction to the psychology of hearing* (5th ed.). Amsterdam: Academic Press.

Mooshammer, C., Hoole, P., & Geumann, A. (2007). Jaw and order. *Language and Speech, 50*(2), 145-176.

Mooshammer, C., Hoole, P., & Kühnert, B. (1995). On loops. *Journal of Phonetics, 23*(1), 3-21.

Mounger, C. (2012). *An Exploration of the Effects of Language on the Orchestral Trombone Sound in France, Germany, and the United States* (Unpublished doctoral dissertation), University of Miami, Coral Gables, FL.

Mukai, S. (1989). Laryngeal movements during wind instruments play. *Nihon Jibiinkoka Gakkai kaiho, 92*(2), 260-270.

Mukai, M. (1992). *Laryngeal movement while playing wind instruments.* Paper presented at the International Symposium on Musical Acoustics (ISMA), Tokyo, Japan.

Müller, E. M., & Brown, W. (1980). Variations in the supraglottal air pressure waveform and their articulatory interpretation. *Speech and Language: Advances in Basic Research and Practice, 4*, 317-389.

Müller, J., Wenning, G. K., Verny, M., McKee, A., Chaudhuri, K. R., Jellinger, K., … Litvan, I. (2001). Progression of dysarthria and dysphagia in postmortem-confirmed parkinsonian disorders. *Archives of Neurology, 58*(2), 259-264.

Munhall, K. G., Löfqvist, A., & Kelso, J. S. (1994). Lip–larynx coordination in speech: Effects of mechanical perturbations to the lower lip. *The Journal of the Acoustical Society of America, 95*(6), 3605-3616.

Munhall, K. G., Ostry, D., & Flanagan, J. (1991). Coordinate spaces in speech planning. *Journal of Phonetics, 19*, 293-307.

Myers, A., Pyle Jr, R. W., Gilbert, J., Campbell, D. M., Chick, J. P., & Logie, S. (2012). Effects of nonlinear sound propagation on the characteristic timbres of brass instruments. *The Journal of the Acoustical Society of America, 131*(1), 678-688.

Napadow, V. J., Kamm, R. D., & Gilbert, R. J. (2002). A biomechanical model of sagittal tongue bending. *Journal of Biomechanical Engineering, 124*(5), 547-556.

Nasir, S. M., & Ostry, D. J. (2006). Somatosensory precision in speech production. *Current Biology, 16*(19), 1918-1923.

Nasir, S. M., & Ostry, D. J. (2008). Speech motor learning in profoundly deaf adults. *Nature Neuroscience, 11*(10), 1217-1222.

Nazari, M. A., Perrier, P., Chabanas, M., & Payan, Y. (2011). Shaping by stiffening: A modeling study for lips. *Motor Control, 15*(1), 141-168.

Needleman, K. (2016, May 13). Thanks, Wai Kit Leung, for the terrible review! [Facebook post on personal page]. Retrieved (2016, July 2 – since removed) from https://www.facebook.com/katherine.needleman

Nichols, R. L., Hanson, F. S., Daehling, W. A., & Hofmann, H. (1971). The Improvement of Brass Instrument Teaching Through the Use of a Profile of the Physical Aspects Involved. Final Report.

Nishio, M., & Niimi, S. (2004). Relationship between speech and swallowing disorders in patients with neuromuscular disease. *Folia Phoniatrica et Logopaedica, 56*(5), 291-304.

Nogita, A., Yamane, N., & Bird, S. (2013). *The Japanese unrounded back vowel /ɯ/ is in fact rounded central/front [ʉ-ʏ]*. Paper presented at Ultrafest VI, Queen Margaret University, Edinburgh.

Norman, L., Chick, J., Campbell, D., Myers, A., & Gilbert, J. (2010). Player Control of'Brassiness' at Intermediate Dynamic Levels in Brass Instruments. *Acta Acustica united with Acustica, 96*(4), 614-621.

Öhman, S. E. (1967). Numerical model of coarticulation. *The Journal of the Acoustical Society of America, 41*(2), 310-320.

O'Shaughnessy, D. (1981). A study of French vowels and consonant durations. *Journal of Phonetics, 9*, 385-406.

Pantev, C., Roberts, L. E., Schulz, M., Engelien, A., & Ross, B. (2001). Timbre-specific enhancement of auditory cortical representations in musicians. *Neuroreport, 12*(1), 169-174.

Pantev, C., Ross, B., Fujioka, T., Trainor, L. j., Schulte, M., & Schulz, M. (2003). Music and Learning-Induced Cortical Plasticity. *Annals of the New York Academy of Sciences, 999*(1), 438-450.

Patel, A. D., & Iversen, J. R. (2003). *Acoustic and perceptual comparison of speech and drum sounds in the north indian tabla tradition: An empirical study of sound symbolism.* Paper presented at the 15th international congress of phonetic sciences (ICPhS), Barcelona, Spain.

Paul, P. M. (2005). The effect of manipulating sound envelope components on identifying wind instrument timbre. *Bulletin of the Council for Research in Music Education*, 17-26.

Payan, Y., & Perrier, P. (1997). Synthesis of VV sequences with a 2D biomechanical tongue model controlled by the Equilibrium Point Hypothesis. *Speech Communication, 22*(2), 185-205.

Penfield, W., & Roberts, L. (1959). *Speech and brain mechanisms*. Princeton, NJ: Princeton University Press.

Penhallurick, R. (2008). Welsh English: Phonology. In B. Kortmann & C. Upton (Eds.), *Varieties of English - The British Isles* (Vol. 1, pp. 105-121). Berlin: Mouton de Gruyter.

Penzl, H. (1975). *Vom Urgermanischen zum Neuhochdeutschen. Eine historische Phonologie (Grundlagen der Germanistik 16)*: Berlin: Erich Schmidt.

Perkell, J. S. (1969). *Physiology of speech production: Results and implications of a quantitative cineradiographic study*. Cambridge, MA: MIT Press.

Perkell, J. S. (1979). On the nature of distinctive features: implications of a preliminary vowel production study. In B. Lindblom & S. Öhman (Eds.), *Frontiers of speech communication research* (pp. 365-380). London: Academic Press.

Perkell, J. S. (2012). Movement goals and feedback and feedforward control mechanisms in speech production. *Journal of Neurolinguistics, 25*(5), 382-407.

Perkell, J., Guenther, F., Lane, H., Marrone, N., Matthies, M. L., Stockmann, E., ... Zandipour, M. (2006). Production and perception of phoneme contrasts covary across speakers. *Speech production: Models, phonetic processes and techniques*, 69-84.

Perkell, J. S., & Klatt, D. H. (1986). *Invariance and variability in speech processes*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Perkell, J., Matthies, M., Lane, H., Guenther, F., Wilhelms-Tricarico, R., Wozniak, J., & Guiod, P. (1997). Speech motor control: Acoustic goals, saturation effects, auditory feedback and internal models. *Speech Communication, 22*(2), 227-250.

Perkell, J. S., Matthies, M. L., Tiede, M., Lane, H., Zandipour, M., Marrone, N., ... Guenther, F. H. (2004). The distinctiveness of speakers' /s/-/S/ contrast is related to their auditory discrimination and use of an articulatory saturation effect. *Journal of Speech, Language, and Hearing Research, 47*, 1259-1269.

Perrier, P. (2006). About speech motor control complexity. In J. Harrington & M. Tabain (Eds.), *Speech production: Models, phonetic processes, and techniques* (pp. 13-25). New York: Psychology Press.

Perrier, P. (2012). Gesture planning integrating knowledge of the motor plant's dynamics: A literature review from motor control and speech motor control. In S. Fuchs, M. Weirich, D. Pape, & P. Perrier (Eds.), *Speech planning and dynamics* (pp. 191-238). Bern: Peter Lang.

Perrier, P., Ostry, D. J., & Laboissière, R. (1996). The equilibrium point hypothesis and its application to speech motor control. *Journal of Speech, Language, and Hearing Research, 39*(2), 365-378.

Perrier, P., & Winkler, R. (2015). Biomechanics of the orofacial motor system: Influence of speaker-specific characteristics on speech production. In S. Fuchs, D. Pape, C. Petrone, & P. Perrier (Eds.), *Individual differences in speech production and perception* (pp. 223-254). Frankfurt: Peter Lang.

Petitti, E. M., & Perrachione, T. K. (2015). A fundamental bias for residue pitch perception in tone language speakers. In The Scottish Consortium for ICPhS 2015 (Eds.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, Scotland.

Pétursson, M. (1971). Etude de la réalisation des consonnes islandaises þ, ð, s, dans la prononciation d’un sujet islandais à partir de la radiocinématographie. *Phonetica, 23*(4), 203-216.

Pizka, H. (1986). *Hornisten-Lexikon*. Kirchheim: Hans Pizka Edition.

Playing under Toscanini: An alumnus remembers. (2014, September). Retrieved from http://www.juilliard.edu/journal/1409/toscanini-fred-fox

Preisler, A. (1993). The influence of spectral composition of complex tones and of musical experience on the perceptibility of virtual pitch. *Perception & Psychophysics, 54*(5), 589-603.

Pribram, K. H. (1977). Hemispheric specialization: Evolution or revolution. *Annals of the New York Academy of Sciences, 299*(1), 18-22.

Purcell, D. W., & Munhall, K. G. (2006a). Adaptive control of vowel formant frequency: Evidence from real-time formant manipulation. *The Journal of the Acoustical Society of America, 120*(2), 966-977.

Purcell, D. W., & Munhall, K. G. (2006b). Compensation following real-time manipulation of formants in isolated vowels. *The Journal of the Acoustical Society of America, 119*(4), 2288-2297.

Pyle Jr, R. W. (1998). The effect of wall materials on the timbre of brass instruments. *The Journal of the Acoustical Society of America, 103*(5), 2834-2834.

Quinque, R. (1982). *atmung stütze ansatz technik für trompete*. Bulle, Switzerland: Editions BIM.

Ramanarayanan, V., Goldstein, L., Byrd, D., & Narayanan, S. S. (2013). An investigation of articulatory setting using real-time magnetic resonance imaging. *The Journal of the Acoustical Society of America, 134*(1), 510-519.

R Development Core Team (2008). R: A language and environment for statistical computing [computer program]. Vienna, Austria: R Foundation for Statistical Computing.

Reinhardt, D. S. (1973). *Encyclopedia of the Pivot System*. Tappan, NY: Charles Colin Publications.

Rosenblum, L. D. (2008). Speech perception as a multimodal phenomenon. *Current Directions in Psychological Science, 17*(6), 405-409.

Rosenkranz, K., Williamon, A., Butler, K., Cordivari, C., Lees, A. J., & Rothwell, J. C. (2005). Pathophysiological differences between musician's dystonia and writer's cramp. *Brain, 128*(4), 918-931.

Rosenthal Archives of the Chicago Symphony Orchestra. (n. d.) Sir Georg Solti. Retrieved from http://cso.org/media/solti100/bio.aspx

Rosset i Llobet, J. (2005). New Tools for the Assessment of the Embouchure's Biomechanics. *Journal of the International Trumpet Guild, 29*(3), 51-53, 81.

Rudy, K., & Yunusova, Y. (2013). The effect of anatomic factors on tongue position variability during consonants. *Journal of Speech, Language, and Hearing Research, 56*(1), 137-149.

Rydell, R., Karlsson, M., Milesson, A., & Schalén, L. (1996). Laryngeal activity during wind instrument playing: video endoscopic documentation. *Logopedics Phonatrics Vocology, 21*(1), 43-48.

Sanders, I., & Mu, L. (2013). A Three-Dimensional Atlas of Human Tongue Muscles. *The Anatomical Record, 296*(7), 1102-1114.

Scavone, G. P., Lefebvre, A., & da Silva, A. R. (2008). Measurement of vocal-tract influence during saxophone performance. *The Journal of the Acoustical Society of America, 123*(4), 2391-2400.

Schlafer, J. (2006). *A Linguistic Analysis of the Use of Vowels to Affect Voicing on the Bassoon* (Unpublished doctoral dissertation), Arizona State University, Tucson, AZ.

Schlossberg, M. (1937). Daily Drills and Technical Studies for Trumpet. New York: J. & F. Hill.

Schmidt, R. A., & Lee, T. (2011). *Motor control and learning: A behavioral emphasis* (5th ed.). Champaign, IL: Human kinetics.

Schneider, P., Sluming, V., Roberts, N., Scherg, M., Goebel, R., Specht, H. J., ... Rupp, A. (2005a). Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference. *Nature Neuroscience, 8*(9), 1241-1247.

Schneider, P., Sluming, V., Roberts, N., Bleeck, S., & Rupp, A. (2005b). Structural, functional, and perceptual differences in Heschl's gyrus and musical instrument preference. *Annals of the New York Academy of Sciences, 1060*(1), 387-394.

Schneider, P., & Wengenroth, M. (2009). The neural basis of individual holistic and spectral sound perception. *Contemporary music review, 28*(3), 315-328.

Schubert, E., Wolfe, J., & Tarnopolsky, A. (2004). *Spectral centroid and timbre in complex, multiple instrumental textures*. Paper presented at the International Conference on Music Perception and Cognition (ICMPC), North Western University, Illinois.

Schumacher, M., Schmoor, C., Plog, A., Schwarzwald, R., Taschner, C., Echternach, M., ... Spahn, C. (2013). Motor functions in trumpet playing—a real-time MRI analysis. *Neuroradiology, 55*(9), 1171-1181.

Scobbie, J. M., Lawson, E., Cowen, S., Cleland, J., & Wrench, A. A. (2011). A common co-ordinate system for mid-sagittal articulatory measurement. *QMU CASL Working Papers WP-20*.

Scobbie, J. M., Stuart-Smith, J., & Lawson, E. (2012). Back to front: a socially-stratified ultrasound tongue imaging study of Scottish English /u/. *Italian Journal of Linguistics, 24*(1), 103-148.

Seither-Preisler, A., Johnson, L., Krumbholz, K., Nobbe, A., Patterson, R., Seither, S., & Lütkenhöner, B. (2007). Tone sequences with conflicting fundamental pitch and timbre changes are heard differently by musicians and nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance, 33*(3), 743.

Shiller, D. M., Ostry, D. J., & Gribble, P. L. (1999). Effects of gravitational load on jaw movements in speech. *The Journal of Neuroscience, 19*(20), 9073-9080.

Silverman, D. (2011). Schwa. In M. van Oostendorp, C. J. Ewen, E. V. Hume, & K. Rice (Eds.), *The Blackwell Companion to Phonology* (Vol. 2, pp. 628-642). Hoboken, NJ: John Wiley & Sons.

Šimáčková, Š., Podlipský, V. J., & Chládková, K. (2012). Czech spoken in Bohemia and Moravia. *Journal of the International Phonetic Association, 42*(2), 225-232.

Simpson, A. P. (2001). Dynamic consequences of differences in male and female vocal tract dimensions. *The Journal of the Acoustical Society of America, 109*, 2153.

Simpson, A. P. (2002). Gender-specific articulatory–acoustic relations in vowel sequences. *Journal of Phonetics, 30*(3), 417-435.

Simpson, A., & Ericsdotter, C. (2007). *Sex-specific differences in f0 and vowel space.* Paper presented at the 16th International Congress of Phonetic Sciences (ICPhS), Saarbrücken, Germany.

Slotine, J.-J. E., & Li, W. (1991). *Applied nonlinear control*. Englewood Cliffs, NJ: Prentice-Hall.

Smith, R. (1986). The effect of material in brass instruments: A review. *Proceedings of the Institute of Acoustics, 8*(1), 91-96.

Smith, T. S. (1971). A phonetic study of the function of the extrinsic tongue muscles. *UCLA Working Papers in Phonetics, 18*.

Sonies, B. C., Shawker, T. H., Hall, T. E., Gerber, L. H., & Leighton, S. B. (1981). Ultrasonic visualization of tongue motion during speech. *The Journal of the Acoustical Society of America, 70*(3), 683-686.

Spahn, C., Richter, B., Pöppe, J., & Echternach, M. (2013). Physiological Insights for Players of Wind Instruments [DVD]. Innsbruck: Helbling.

Speer, D. (1973). *Ungarischer oder Dacianischer Simplicissimus* (M. Szyrocki & K. Gajek Eds.). Vienna: Österr. Bundesverl. für Unterricht, Wiss. und Kunst. (Original work published 1683)

Spiegelberg, S. C. (2002). *The psychoacoustics of musical articulation* (Unpublished doctoral dissertation), University of Rochester, Rochester, NY.

Šram, F., & Švec, J. G. (2000). Die Tonerzeugung beim Spielen von Blasinstrumenten. In J. Pahn (Ed.), *Sprache und Musik: Beiträge der 71. Jahrestagung der Deutschen Gesellschaft für Sprach-und Stimmheilkunde e. V.* (pp. 155-159). Berlin, Germany.

Stauffer, D. W. (1954). Intonation deficiencies of wind instruments. Washington: The Catholic University of America Press.

Stavness, I., Gick, B., Derrick, D., & Fels, S. (2012). Biomechanical modeling of English /r/ variants. *The Journal of the Acoustical Society of America, 131*(5), EL355-EL360.

Stavness, I., Lloyd, J. E., & Fels, S. (2012). Automatic prediction of tongue muscle activations using a finite element model. *Journal of Biomechanics, 45*(16), 2841-2848.

Steenstrup, K. (2007). *Teaching brass*. Aarhus, Denmark: Aarhus University Press.

Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics, 17*, 3-45.

Stevens, K. N. (2002). Toward a model for lexical access based on acoustic landmarks and distinctive features. *The Journal of the Acoustical Society of America, 111*(4), 1872-1891.

Stevens, K. N., & Keyser, S. J. (2010). Quantal theory, enhancement and overlap. *Journal of Phonetics, 38*(1), 10-19.

Stevenson, A. J., Chiu, C., Maslovat, D., Chua, R., Gick, B., Blouin, J.-S., & Franks, I. M. (2014). Cortical involvement in the StartReact effect. *Neuroscience, 269*, 21-34.

Stewart, M. D. (Ed.). (1987). *Arnold Jacobs: the legacy of a master: the personal and pedagogical recollections of thirty-one of his colleagues, students, and friends*. Northfield, IL: The Instrumentalist Company.

Stolar, S., & Gick, B. (2013). An index for quantifying tongue curvature. *Canadian Acoustics, 41*(1).

Stone, M. (1990). A three-dimensional model of tongue movement based on ultrasound and x-ray microbeam data. *The Journal of the Acoustical Society of America, 87*, 2207.

Stone, M. (2005). A guide to analysing tongue motion from ultrasound images. *Clinical Linguistics & Phonetics, 19*(6-7), 455-501.

Stone, M., & Davis, E. P. (1995). A head and transducer support system for making ultrasound images of tongue/jaw movement. *The Journal of the Acoustical Society of America, 98*(6), 3107-3112.

Stone, M., Epstein, M., & Iskarous, K. (2004). Functional segments in tongue movement. *Clinical Linguistics & Phonetics, 18*(6-8), 507-521.

Stone, M., Epstein, M., Li, M., & Kambhamettu, C. (2006). Predicting 3D tongue shapes from midsagittal contours. *Speech production: Models, phonetic processes, and techniques*, 315-330.

Stone, M., & Lundberg, A. (1996). Three-dimensional tongue surface shapes of English consonants and vowels. *The Journal of the Acoustical Society of America, 99*, 3728.

Stone, M., Shawker, T. H., Talbot, T. L., & Rich, A. H. (1988). Cross-sectional tongue shape during the production of vowels. *The Journal of the Acoustical Society of America, 83*, 1586.

Stone, M., & Vatikiotis-Bateson, E. (1995). Trade-offs in tongue, jaw, and palate contributions to speech production. *Journal of Phonetics, 23*(1–2), 81-100.

Stuart-Smith, J., Macdonald, R., Jose, B., & Evers, L. (2015). *Considering the evidence for real-time change in a standard variety: acoustic vowel quality in Scottish Standard English in Glasgow*. Poster presented at the 10th UK Language Variation and Change Conference (UKLVC), York, England.

Studdert-Kennedy, M., & Goldstein, L. (2003). Launching language: The gestural origin of discrete infinity. *Studies in the Evolution of Language, 3*, 235-254.

Strong, L. H. (1956). Muscle fibers of the tongue functional in consonant production. *The Anatomical Record, 126*(1), 61-79.

Sumbre, G., Fiorito, G., Flash, T., & Hochner, B. (2006). Octopuses use a human-like strategy to control precise point-to-point arm movements. *Current Biology, 16*(8), 767-772.

Švec, J. G., & Schutte, H. K. (1996). Videokymography: high-speed line scanning of vocal fold vibration. *Journal of Voice, 10*(2), 201-205.

Sweet, H. (1890). *A primer of phonetics*. Oxford: Clarendon Press.

Takemoto, H. (2001). Morphological analyses of the human tongue musculature for three-dimensional modeling. *Journal of Speech, Language, and Hearing Research, 44*(1), 95-107.

Tarnopolsky, A., Fletcher, N., Hollenberg, L., Lange, B., Smith, J., & Wolfe, J. (2005). Acoustics: the vocal tract and the sound of a didgeridoo. *Nature, 436*(7047), 39-39.

Tarnopolsky, A. Z., Fletcher, N. H., Hollenberg, L. C., Lange, B. D., Smith, J., & Wolfe, J. (2006). Vocal tract resonances and the sound of the Australian didjeridu (yidaki) I. Experiment. *The Journal of the Acoustical Society of America, 119*(2), 1194-1204.

Tarr, E. (1988). *The trumpet* (S. E. Plank & E. Tarr, Trans.). Portland, OR: Amadeus Press.

Tarr, E. H., & Dickey, B. (2007). *Articulation in early wind music: a source book with commentary* (Vol. 8). Winterthur: Amadeus.

Taumoefolau, M. (2002). Stress in Tongan. *MIT Working papers in Linguistics (8th Meeting of the Austronesian Formal Linguistics Association (AFLA)), 44*.

Taylor, R. (1996). Old French. In T. J. McGee, A. G. Rigg, & D. N. Klausner (Eds.), *Singing Early Music: The pronunciation of European languages in the late Middle Ages and Renaissance* (pp. 65-89). Bloomington, IN: Indiana University Press.

Thayer, R. C. (1974). The effect of the attack transient on aural recognition of instrumental timbres. *Psychology of Music, 2*(1), 39-52.

Thibeault, M., Ménard, L., Baum, S. R., Richard, G., & McFarland, D. H. (2011). Articulatory and acoustic adaptation to palatal perturbation. *The Journal of the Acoustical Society of America, 129*(4), 2112-2120.

Tiede, M. K., Masaki, S., Wakumoto, M., & Vatikiotis-Bateson, E. (1997). Magnetometer observation of articulation in sitting and supine conditions. *The Journal of the Acoustical Society of America, 102*(5), 3166-3166.

Tiede, M., & Whalen, D. (2015). *GetContours: an interactive tongue surface extraction tool*. Paper presented at Ultrafest VII, Hong Kong, China.

Titze, I. R. (2008). Nonlinear source–filter coupling in phonation: Theory. *The Journal of the Acoustical Society of America, 123*(5), 2733-2749.

Todorov, E., & Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neuroscience, 5*(11), 1226-1235.

Torres, E. H. (2012). *Flute articulation pedagogy: The effect of language-specific consonant pronunciation on a flutist's articulation within the French and English languages* (Unpublished doctoral dissertation), The Ohio State University, Columbus, OH.

Torres-Oviedo, G., & Ting, L. H. (2010). Subject-specific muscle synergies in human balance control are consistent across different biomechanical contexts. *Journal of Neurophysiology, 103*(6), 3084-3098.

Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes, 26*(7), 952-981.

Traser, L., Burdumy, M., Richter, B., Vicari, M., & Echternach, M. (2013). The effect of supine and upright position on vocal tract configurations during singing—a comparative study in professional tenors. *Journal of Voice, 27*(2), 141-148.

Tremblay, S., Houle, G., & Ostry, D. J. (2008). Specificity of speech motor learning. *The Journal of Neuroscience, 28*(10), 2426-2434.

Trulsson, M., & Essick, G. K. (1997). Low-threshold mechanoreceptive afferents in the human lingual nerve. *Journal of Neurophysiology, 77*(2), 737-748.

Tu'inukuafe, E. (1992). *A simplified dictionary of modern Tongan*. Auckland: Polynesian Press.

Tuller, B., Harris, K. S., & Gross, B. (1981). Electromyographic study of the jaw muscles during speech. *Journal of Phonetics, 9*, 175-188.

Turnbull, D. R. (2001). *An analysis, clarification, and revaluation of Donald Reinhardt's pivot system for brass instruments* (Unpublished doctoral dissertation), Arizona State University, Tempe, AZ.

Turvey, M. T. (1977). Preliminaries to a theory of action with reference to vision. In R. E. Shaw & J. Bransford (Eds.), *Perceiving, acting and knowing: Toward an ecological psychology* (pp. 211-265). Hillsdale, NJ: Erlbaum.

Turvey, M. T., Shaw, R. E., & Mace, W. (1978). Issues in the theory of action: Degrees of freedom, coordinative structures and coalitions. *Attention and performance VII*, 557-595.

Valls-Solé, J., Rothwell, J. C., Goulart, F., Cossu, G., & Munoz, E. (1999). Patterned ballistic movements triggered by a startle in healthy humans. *The Journal of Physiology, 516*(3), 931-938.

Vatikiotis-Bateson, E., & Kelso, J. (1990). Linguistic structure and articulatory dynamics: A cross language study. *Haskins Laboratories Status Report on Speech Research*, 67-94.

Vietor, W. (1884). *Elemente der Phonetik des Deutschen, Englischen und Französischen*. Leipzig: Reisland.

Villacorta, V. M., Perkell, J. S., & Guenther, F. H. (2007). Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. *The Journal of the Acoustical Society of America, 122*(4), 2306-2319.

Vogt, F., Lloyd, J., Buchaillard, S., Perrier, P., Chabanas, M., Payan, Y., & Fels, S. (2006). *An efficient biomechanical tongue model for speech research.* Paper presented at the Proceedings of the 7th International Seminar on Speech Production.

Vorperian, H. K., Kent, R. D., Lindstrom, M. J., Kalina, C. M., Gentry, L. R., & Yandell, B. S. (2005). Development of vocal tract length during early childhood: A magnetic resonance imaging study. *The Journal of the Acoustical Society of America, 117*(1), 338-350.

Wadman, W. J., Van Der Gon, J. D., Geuze, R. H., & Mol, C. (1979). Control of fast goal-directed arm movements. *Journal of Human Movement Studies, 5*(1), 3-17.

Wagner, P., Trouvain, J., & Zimmerer, F. (2015). In defense of stylistic diversity in speech research. *Journal of Phonetics, 48*, 1-12.

Wallis, J. (1972). *Grammatica linguae anglicanae* (J. A. Kemp, Trans. J. A. Kemp Ed.). London: Longman. (Original work published 1653)

Weast, R. D. (1963). A stroboscopic analysis of lip function. *The Instrumentalist, 23*, 44-48.

Weber, K. (1978). Die "deutsche" Posaune. *Das Orchester, 7/8*, 566-570.

Wein, B., Böckler, R., Sram, F., Axmann, K., & Etlik, M. (1989). *Ultrasonic animations of tongue shapes during playing wind-instruments.* Paper presented at the 21st Congress of the International Association of Logopedics (IALP), Prague, Czechoslovakia.

Weirich, M., & Simpson, A. P. (2014). *Articulatory vowel spaces of male and female speakers.* Paper presented at the Proceedings of the 10th International Seminar on Speech Production (ISSP).

Wells, J. C. (1982). *Accents of English.* Cambridge: Cambridge University Press.

Wells, J. C. (1995). *Computer-coding the IPA: A proposed extension of SAMPA*. Revised draft.

Westbury, J. R., & Keating, P. A. (1986). On the naturalness of stop consonant voicing. *Journal of Linguistics, 22*(01), 145-166.

Whalen, D. H., Iskarous, K., Tiede, M. K., Ostry, D. J., Lehnert-LeHouillier, H., Vatikiotis-Bateson, E., & Hailey, D. S. (2005). The Haskins optically corrected ultrasound system (HOCUS). *Journal of Speech, Language & Hearing Research, 48*(3), 543-553.

Whiteside, S. P. (2001). Sex-specific fundamental and formant frequency patterns in a cross-sectional study. *The Journal of the Acoustical Society of America, 110*(1), 464-478.

Wickham, H., Chang, W., & RStudio. (2016). Package 'ggplot2': An Implementation of the Grammar of Graphics [Computer program]. Version 2.1.

Wiggins, G. C., & Storey, P. (2010). *Rapid frame-rate MR acquisitions to reveal mechanisms of circular breathing and sound production in the Australian Aboriginal didgeridoo.* Paper presented at the Annual meeting of the International Society for Magnetic Resonance in Medicine (ISMRM), Stockholm, Sweden.

Wilson, I. L. (2006). *Articulatory settings of French and English monolingual and bilingual speakers* (Unpublished doctoral dissertation), University of British Columbia, Vancouver, Canada.

Wilson, I., & Gick, B. (2014). Bilinguals use language-specific articulatory settings. *Journal of Speech, Language, and Hearing Research, 57*(2), 1-13.

Wilson, I., Horiguchi, N., & Gick, B. (2007). *Japanese articulatory setting: the tongue, lips and jaw*. Paper presented at UltraFest IV, New York University.

Wilson, T. D. (1996). *The measured upstream impedance for clarinet performance and its role in sound production* (Unpublished doctoral dissertation), University of Washington, Seattle, WA.

Winckel, F. (1967). *Music, sound and sensation: A modern exposition* (T. Binkley, Trans.). New York: Dover Publications.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). *ELAN: A professional framework for multimodality research.* Paper presented at the 5th International Conference on Language Resources and Evaluation, Genoa, Italy.

Wolfe, J., Almeida, A., Chen, J. M., George, D., Hanna, N., & Smith, J. (2013). *The player–wind instrument interaction*. Paper presented at the Stockholm Music Acoustics Conference (SMAC), Stockholm, Sweden.

Wolfe, J., Chen, J.-M., & Smith, J. (2010). *The acoustics of wind instruments—and of the musicians who play them*. Paper presented at the International Congress on Acoustics (ICA), Sydney, Australia.

Wolfe, J., Fletcher, N. H., & Smith, J. (2015). Interactions between wind instruments and their players. *Acta Acustica united with Acustica*(101), 211-223.

Wolfe, J., Garnier, M., & Smith, J. (2009). Vocal tract resonances in speech, singing, and playing musical instruments. *Human Frontier Science Program Journal, 3*(1), 6-23.

Wolfe, J., & Smith, J. (2008). *Acoustical coupling between lip valves and vocal folds.* Paper presented at Acoustics Australia.

Wolfe, J., Tarnopolsky, A., Fletcher, N., Hollenberg, L., & Smith, J. (2003). *Some effects of the player's vocal tract and tongue on wind instrument sound*. Paper presented at the Stockholm Music Acoustics Conference (SMAC), Stockholm, Sweden.

Wong, P. C. (2002). Hemispheric specialization of linguistic pitch patterns. *Brain research bulletin, 59*(2), 83-95.

Wong, P. C., Warrier, C. M., Penhune, V. B., Roy, A. K., Sadehh, A., Parrish, T. B., & Zatorre, R. J. (2008). Volume of left Heschl's gyrus and linguistic pitch learning. *Cerebral Cortex, 18*(4), 828-836.

Worman, W. E. (1971). *Self-Sustained Nonlinear Oscillations of Medium Amplitude in Clarinet-Like Systems* (Unpublished doctoral dissertation), Case Western Reserve University, Cleveland, OH.

Wrench, A. A., Cleland, J., & Scobbie, J. M. (2011). *An ultrasound protocol for comparing tongue contours: Upright vs. supine*. Paper presented at the 17th International Congress of Phonetic Sciences (ICPhS), Hong Kong, China.

Wright, H., & Campbell, D. (1998). *The influence of the mouthpiece on the timbre of cup-mouthpiece wind instruments*. Paper presented at the International Symposium on Musical Acoustics (ISMA), Leavenworth, WA.

Yamane-Tanaka, N., Gick, B., & Bird, S. (2004). Bimodal schwa: Evidence from acoustic measurements. *The Journal of the Acoustical Society of America, 115*(5), 2542-2542.

Yoshikawa, S. (1995). Acoustical behavior of brass player's lips. *The Journal of the Acoustical Society of America, 97*(3), 1929.

Yoshikawa, S. (1998). Vibration labiale et contrôle du souffle chez les joueurs de cuivres. *Médecine des Arts, 26*, 22-26.

Yoshikawa, S., & Muto, Y. (2003). Lip-wave generation in horn players and the estimation of lip-tissue elasticity. *Acta Acustica united with Acustica, 89*(1), 145-162.

Young, S., Evermann, G., Gales, M., Hain, T., Kershaw, D., Liu, X. A., ... Povey, D. (2006). *The HTK book (for HTK version 3.4)*: Cambridge University Engineering Department.

Zharkova, N. (2013a). A normative-speaker validation study of two indices developed to quantify tongue dorsum activity from midsagittal tongue shapes. *Clinical Linguistics & Phonetics, 27*(6-7), 484-496.

Zharkova, N. (2013b). Using ultrasound to quantify tongue shape and movement characteristics. *The Cleft Palate-Craniofacial Journal, 50*(1), 76-81.

Ziegler, W. (2002). Task-related factors in oral motor control: Speech and oral diadochokinesis in dysarthria and apraxia of speech. *Brain and language, 80*(3), 556-575.

Ziegler, W. (2003). Review: Speech motor control is task-specific: Evidence from dysarthria and apraxia of speech. *Aphasiology, 17*(1), 3-36.

Ziegler, W. (2006). Distinctions between speech and non-speech motor control: A neurophonetic view. In J. Harrington & M. Tabain (Eds.), *Speech production: Models, phonetic processes, and techniques* (pp. 41-54). New York: Psychology Press.

Zielke, A. (2010). *Zungensonographie und Gesichts-Hals-Motorik beim Spielen von Blasinstrumenten* (Unpublished doctoral dissertation), Heinrich-Heine Universität, Düsseldorf.

Zimmerer, F. (2009). *Reduction in natural speech* (Unpublished doctoral dissertation), Johann Wolfgang Goethe-Universität, Frankfurt am Main.

Zsaitsits, M. (2012). *Zur Zunge beim Spiel von Blechblasinstrumenten* (Unpublished MA dissertation), Universität für Musik und darstellende Kunst Wien, Wien.

Appendices

Appendix A: Online questionnaire and corresponding Human Ethics approval

Metainformation:

Title and small print appearing on top of each slide:
The Influence of First Language upon playing a Brass Instrument

*Answers with a * were required to advance to the next page.*

-- *(marks page breaks)*

1. Information for Participants of Online Questionnaire

Matthias Heyne
Ph.D. project at the University of Canterbury, Christchurch, New Zealand
Supervised by Prof. Dr. Jennifer Hay

Thank you for offering to participate in my research project on the influence of First Language upon playing a brass instrument. In my Ph.D. project at the University of Canterbury, I will investigate whether one's First Language/s influence/s the way one plays a brass instrument and this pilot study is meant to collect comparable data from players with different language backgrounds to help identify relevant areas of influence.

I am myself a bass trombonist with degrees in orchestral music and jazz and for this reason I am interested in the question why brass players from different national backgrounds sound differently, while, at the same time, I have become interested in Linguistics as part of studies towards becoming a teacher for Music and English in High School.

In the main part of the proposed research project, I will use ultrasound to record brass players' tongue positioning in order to obtain empirically documented data on what players from diverse language groups may do differently which leads to perceivable differences in playing.

Your involvement in this pilot study will consist of filling out a questionnaire that includes questions on your personal background, learning and playing your brass instrument, as well as national schools/styles of playing and how you perceive them, if at all. The questionnaire comprises a total of 25 questions and takes about 20 minutes to compete. Only questions marked with a * are required.

Participation is voluntary and you have the right to withdraw at any stage by clicking "Exit"; no entries will be saved in this case. If you would merely like to suspend filling out the questionnaire, you can do that by clicking on the button "Save and continue later and you will be shown a link that you need to save in order to continue later on. Any questionnaire that is incomplete in the sense that one or more required questions (marked by an asterisk) have not been answered will be excluded from the results obtained through this research. Should you have completed the questionnaire and later decide to withdraw your participation, please write me an email at matthias.heyne@pg.canterbury.ca.nz and I will remove your dataset if it can be properly identified.

The results obtained through this questionnaire may be published, but you may be assured of the complete confidentiality of the data gathered in this investigation: your identity will not be made public without your prior consent. To ensure anonymity and confidentiality, the data gathered through this online questionnaire will be stored in a secure place and password-protected. I have made sure that www.esurveyspro.com's privacy policy as well as their terms and conditions are in accordance with UC Human Ethics guidelines and this questionnaire has been reviewed and approved by the University of Canterbury Human Ethics Committee. Participants should address any complaints to The Chair, Human Ethics Committee, University of Canterbury, Private Bag 4800, Christchurch, New Zealand (human-ethics@canterbury.ac.nz).

Access to the data will be restricted to the researcher and others involved in the research project at the University of Canterbury, i.e. the supervisor and potential research associates.

The resulting Ph.D. thesis is a public document and will be available through the UC Library. You may receive a copy of my Ph.D. thesis by contacting me at the conclusion of the project (duration approximately 2-3 years) and/or receive a summary of the results of the questionnaire (within the next half year) by sending me an email at Mattes.Heyne@gmx.de. Please note that this email address is different from the one above (matthias.heyne@pg.canterbury.ac.nz), which is my email address at the University of Canterbury that will terminate once I have completed my Ph.D. project.

This project is being carried out as a pilot study for my PhD research project at the University of Canterbury, Christchurch, New Zealand, under the supervision of Prof. Dr. Jennifer Hay, who can be contacted at jen.hay@canterbury.ac.nz. Dr. Hay will be pleased to discuss any concerns you may have about participation in the project.

If you agree to participate in the study, you are asked to carefully read through each of the statements listed on the next slide and by clicking "Next" at the bottom of that page, you will express your consent to those statements.

----

2. Consent Form for Pilot Study: Online Questionnaire

I have carefully read through the above information about this project and will contact the researcher at matthias.heyne@pg.canterbury.ca.nz if there are any questions that I need to have answered before giving my consent and filling out the following questionnaire.

I understand what is required of me if I agree to take part in the research.

I understand that participation is voluntary and I may withdraw at any time without penalty. Withdrawal of participation will also include the withdrawal of any information I have provided should this remain practically achievable.

I understand that any information or opinions I provide will be kept confidential to the researcher and his research associates and that any published or reported results will not identify the participants. I understand that the resulting thesis is a public document and will be available through the UC Library.

I understand that all data collected for the study will be kept in password-protected electronic form.

I understand the risks associated with taking part and how they will be managed.

I understand that I am able to receive a report on the findings of the study by contacting the researcher at the conclusion of the Ph.D. project and/or close of the questionnaire.

I understand that I can contact the researcher (matthias.heyne@pg.canterbury.ac.nz) or supervisor (Prof. Dr. Jennifer Hay, jen.hay@canterbury.ac.nz) for further information. If I have any complaints, I can contact the Chair of the University of Canterbury Human Ethics Committee, Private Bag 4800, Christchurch, New Zealand (human-ethics@canterbury.ac.nz).

By clicking “Next” below, I agree to participate in this research project.

----

3. Personal Information

1. Please provide the following information:*

Age (in years):	____________________
Sex (m/f):	____________________
Town:	____________________
Country of Residence:	____________________
Nationality (add Ethnicity if applicable):	____________________

2. Contact Details (optional):

First Name (enter Middle Names here if applicable):	____________________
Last Name:	____________________
Email address:	____________________
Phone Number (including country code):	____________________

----

4. Brass playing

1. Which brass instrument(s) do/did you play?*
drop down list:
Trumpet
French Horn
Trombone
Tuba

Please use box below if playing an instrument that is not listed or to enter multiple instruments.

____

2. For how long have you played/did you play this/these instrument(s) (in years)? (If playing more than one brass instrument please provide total number of years)*

3. What is your proficiency level regarding the instrument(s) you entered above? (If playing more than one brass instrument please tick only highest level.)*

Professional Semi-professional Amateur

4. If you ticked professional or semi-professional above, with what kind of work do you usually earn your money?

Fixed contract with an Orchestra, Big Band etc.
Music Teaching
Freelance
Other (Please Specify):

____

5. How often do you play your instrument(s) on average (in hours per week, combine all brass instruments if applicable)?*

6. I mostly play in the following contexts:*

Symphony Orchestra
Wind Orchestra/Band
Brass Band
Chamber Music
Solo
Jazz – section player
Jazz – improviser
Other (Please specify):

____

----

5. Language(s)

1. What is your First Language? (If you have more than one mother tongue, please also list other language/s)*
____________________ *(lots of space)*

2. What Second Languages do you speak? (Please order Second Languages according to proficiency level, starting with the Second Language you are most proficient in.)

1 ____________________
2 ____________________
3 ____________________
4 ____________________
5 ____________________

3. For how long have you learned/used these languages? (in years, please order like above)

1 ____________________
2 ____________________
3 ____________________
4 ____________________
5 ____________________

4. How proficient do you consider yourself in these languages? (please order like above)

	Elementary Proficiency	Intermediate Proficiency	Advanced Proficiency	Professional Proficiency	Native-like Proficiency
1					
2					
3					
4					
5					

(boxes to be ticked)

5. Is your dialect/accent when speaking any of the languages mentioned above (First and/or Second Languages) typically associated with a particular region? If so, please provide details.
____________________ *(lots of space)*

6. Have you ever been diagnosed with any speech and/or hearing difficulties? If so, please describe.
____________________ *(lots of space)*

----

6. Brass Playing ↔ Language(s) ↔ National Schools/Styles of Brass Playing

1. In which country/countries did you learn to play your instrument(s)?*
____________________ *(lots of space)*

2. Do you believe that national schools/styles exist regarding (classical) music in general or specific instruments/groups of instruments?*
yes probably undecided probably not no

3. Which school(s)/style(s) did you learn to play? In which school(s)/style(s) are you playing now? (Provide all schools/styles.)*
____________________ *(lots of space)*

4. Which national schools/styles do you think exist? Have you noticed any current (global) developments regarding this phenomenon? Please explain.
____________________ *(lots of space)*

5. Do you think that a person's first language and/or other acquired languages have some kind of influence upon playing brass instruments?*

no influence whatsoever
limited influence
influences playing somewhat
clear influence
much influence
(arranged horizontally in online questionnaire)

6. If you answered affirmatively above: in which areas do you feel that this influence takes place?

Articulation
Sound/Timbre
Breathing
Rhythm
Vibrato
other (Please Specify):
____

7. Where have you noticed this influence?

in my own playing
in others' playing: students that I have taught
in others' playing: in an ensemble/orchestra
other (Please Specify):

8. If you ticked any of the boxes on the question above: Please provide further details.
____________________ *(lots of space)*

9. Try to explain where these influences come from?
____________________ *(lots of space)*

10. Have you ever (been) taught to reproduce or imagine specific speech sound configurations in your playing, and if so which ones?*
___________________ *(lots of space)*

11. Which factor do you think is more influential in affecting brass playing, one's First Language/s (and possibly Second Languages) or playing styles (nationals schools etc.)?*

Language has most influence
Language is somewhat more influential
Undecided
Style is somewhat more influential
Style has most influence
(arranged horizontally in online questionnaire)

----

7. You're almost finished now!

1. Room for further comments:
___________________ *(lots of space)*

----

8. Thank you!

Please contact me at matthias.heyne@pg.canterbury.ac.nz if you should have any questions or comments regarding this questionnaire.
(Permanent email address: *removed from thesis copy*)

Many thanks for taking the time to fill out my questionnaire!

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

HUMAN ETHICS COMMITTEE

Secretary, Lynda Griffioen
Email: human-ethics@canterbury.ac.nz

Ref: HEC 2013/39/LR-PS

23 October 2013

Matthias Heyne
Department of Linguistics
UNIVERSITY OF CANTERBURY

Dear Matthias

Thank you for forwarding to the Human Ethics Committee a copy of the low risk application you have recently made for your research proposal "The influence of first language upon playing a brass instrument pilot study: online questionnaire".

I am pleased to advise that this application has been reviewed and I confirm support of the Department's approval for this project.

Please note that this approval is subject to the incorporation of the amendments you have provided in your email of 22 October 2013.

With best wishes for your project.

Yours sincerely

Lindsey MacDonald
Chair, Human Ethics Committee

University of Canterbury Private Bag 4800, Christchurch 8140, New Zealand. www.canterbury.ac.nz

E S

Appendix B: Materials relating to the ultrasound study, including Human Ethics approval

Tongan wordlist

ā - ā'a - āfei - āata

kī – kīkīvoi - kītaki

teke - tele'a – temipale - tenisi

'aneafi - 'aonga - 'apele - 'angelo

ki – kia - ki'i – kikī - kili

lā – lālanga – lākanga - lāpisi

same – sai - sa'ia - salati

timi – tina – tipiloma - tisi

lēpati – lētisi – lēsisita - lēvolo

kuonga – kupenga – kutu - kuusi

longoā'a – loloto - lose

tafa – taimi – takele - talamu

'a - 'a'ahi - 'afa - 'ahi

konga – koniseti – koosi - kopa

'uanga - 'ufi'ufi - 'uha - 'uhila

laaka - la'e – lafalafa - lahi

sāliote - Sāpate – sātine - sāvolo

luo – lusa – lupe – luva - lū

li'aki – lifi – lili – lilifa - lenini

lea - lea'aki - le'ei - lei

kosipeli – kote – kouna - kovi

ta – tatā - ta'ahine - ta'efe'unga

kohikohi – koko – kolisi - koma

tamaiki – tangakalī – taeakalami - tao

lēisi – lēkooti – lēlue - lēmani

iku - iku'i – ila - ilifia

sea - se'e – seifi - sekisofoni

ongo – onioni – ono - onopooni

lisi – lita – liukava - livi

neivi – nekativi – nenefu - neongo

tī – tīpī – tīkoni – tīpota - tīsolo

teau - te'elango – teemi - teka

tāheu – tāketi – tānaki - tānolo

tepi – tesi – tete - teunga

tapu – tavilo – tasilo - tau

alanga - alame'a – aleapau - angiiki

'āua! - 'āvea - 'āunofu

tia – tikatele – tikisinale - tikite

ua – uafu – uaine - ueita

tūkia'anga – Tūsite - tūtū'anga

kisikisi – kita - kiū

lepa – lepi – letiō - levu

'ene - 'epani - 'esitimete - 'etita

ofi - ofongi – oka - ola

ngū – ngūfeke – ngūngū - ngūnifo

lātisi – lāulau - lāvaki

tiupi – tiuta - tōpiki - tōtō

tuna - tunga'iva'e – tupenu - tutu

nuiki – nunu - nunu'a – nusi - nusipepa

solova – sone - suka

oli – oloveti – oma - ono

tō – tōatu – tōmui - tōngofua

selilī – semipione – sengai

tā – tāta – tāfakatātā - tāfue

ko – koa - koe'uhi - kofi

Sepitema - sesele – seti – sevāniti

kūnima – kūtu - ngingila

sasipani – satelaite – Saute - savea

sipika – siteisi – Siulai - sivi

afā – aka – ako - ala

ni – nifo - ni'ihi - nima
lingitoto – limu – lingi - lipooti
niti – niusi - līpine - lītaula
tui – tukituki – tulama - tumutumu
keukeu – kenitaki - kete
uho – uku – ula - ulofi
lī - lī'angaveve – līekina
kata – kamata – kananga – kanisā
nonga – noniseni – nonofo - no’'ounu
'i ai - 'ia - 'ikai - 'ile'ila - 'ita
keke - kele'a – kemikale - kepi
lotu - louhi’inima - lova
nge'esi - nge'esinima – ngeia - ngeli
'āmio - 'akilotoa - 'ānaki - 'āsili
'ata'atā - 'aupito - 'avalisi
āma – āāpe – ī - ē - ēsei
'ea - 'efi - 'Eiki - 'eve'eva
kōmiti – kōpate – kōpila - Kōvana
loea - lofa - loka'i - lokiako
ngā'ito – ngāngaahi – ngāue - ngāngāue
ninimo – nipi – niu - niumōnia
kapa – kasa – kau - kava
ina – inu – ipu - ivi
ngutu – ngutungutu - nugtumālie
tēkina – tēnoa – tēnipi - tēniti
to'ohema – tosi - totitoti
'umata - 'univēsiti - 'uto
namu – nanamu - napikeni - nauna
lefu – lekomeni – lele - le'o
ka – kafu – kai – kakā – kalama
ke – kehekehe – keipolo
'inisēkite - 'iote - 'isite - 'itāni
tōkia – tōkovi – tōnoa - tō’ohi

lua – lue – lula – lulu - luma'i
sinamoni – singi – sio
'Okatopa - 'oke - 'olive - 'one'one
tominō – Tonga – toni - tou
tu'u – nāunau - natimeki - natule
kē – kēmisi - kēnolo
na - na'a - nafa - nailoni
ngoto’umu – ngoue - 'ōvava
laifolo – laione - la'itā - lakanga
toafa - to'a – toe - tofuā'a
kō atu – kōfekofe – kōlila - kōmiki
engaenga – epu - eve
sēpuni – Sēsu - sētesi
amohi – anga – angalelei - ava
'ukulele - 'uli'uli - 'ulupoko
tohi – toi - to'i – toke - tolofini
lalata - lami - langa’ulu - lanu
aake – aami – aati - afi
tāpolo – tāpuni – tātā - tātātau
ngulu - ngulungulu
kulī – kumā - kumukumu
'eiki - 'ekueta - 'elelo - 'emipola
tu'a – tuaine – tufakanga - tufu
'oanga - 'o - 'ofa - 'ofisa - 'ohofi
sā – sākalameta – sākisi
ifi - ifo – ihu – ika - iviivi
noa – Noate - nofo ā - nofohili
lao – lapa – lau - lavea
loi - lo'imata – lole - lomi
kuata – kuhū – kui - kuku
'opositi - 'Oseni 'Initia - 'otomētiki
sia – siesi - sifi po'uli - si'i
samani – sanipepa – saoa - sapa
lōmuku – lōoa – lōpini - lōua

ngao – ngata – ngau - nga'unu
kā – kāingalotu – kaka - kākaa'i
kālepi – kāloti – kāmeli - kānita
uma – umi – uoti – utu - u'u
ne - neave – neesi - ne'ine'i
'akapulu - 'alafapeti - 'amatua
lō – lōketi – lōlahi – lōloa - lōmaki'i
'ū - 'ūkuma - 'ūpē
ngai'i – ngakongako – ngalingali
sofa – soka – sokaleti
ō - ō'i - ū'a - ūū
efiafi – efuefu – ena - eni

sīlongo - Sīsu – sīnaki - sītu'a
siipi – sikā – sila - sili
kāponi – kāsolo – kātaki - kātoa
tēpile – tētē – tēvolo - ouau
nota – noti – nouti - novesia
kilouati - kimu'a - kinitakāteni
ngaahi – ngafa - ngaholo
ngofua – ngoto - ngongohe
nō – nōpele – nōvolo - Nōvema
sēini – sēlue – seniti
'ilongofua - 'imisi - 'Ingilisi

NZE wordlist

harbinger winger singing longing
zed zealot Zack
chemist kennel kettle captain
bird curt burn bat
ticket tissue tiff tin timid
nun nullify nothing numb
shah shard sharp shark
turret tertiary turpentine
lamentable lascivious lagoon
dummy done duck dusk
leaving leech legal leash
garment Ghana gasp guard
shaft shag shall shack
equation version transfusion
nativity Namibia Napoleon narration
nasality
locker long loss recorder
gondola gosh golf
weather which whether witch
dawn daughter daunting
cabin cackle café cascade
heed hood hid heard
tall Taurus toggle Todd
vision abrasion Asian allusion
hospital dirty turtle sort
safari sagacious tibia tip tick
semitic sentence savage sash
turban turner turkey turf
gut gutter gun
double dozen duffel Dutch dungeon
goon Google gooseberry
sherd shirk shirt shirtless
tattoo torrential tequila terrific
shisha she Sheba sheath
second scepter seduce cell
hod horde city
nerve nurse nervous nurture
collision cohesion provisional
Dakota dominion denominator
debauchery
delta desert despair
silly sit soon suit
litter ligament liminal
sin sister cynical sip
search serpent surge certainly
curly curb curve curfew
preservative observance exert
abandon scandalous abidance
academy
courier kung-fu coup
bet bed beck beg
gull gum government gust
shut shuffle shudder shush
later label lamp
exult result resumption presumptive
December decent decoy deed
navigation nephew nest nettle
good goods gourmet goodness
fitted better later butter
deposit tendon accident expendable
zeal zebra tenacious
beat bead beak bean
liver listen linen lip
tonight towards tomorrow
ludicrous loom Lucas lucent
sociable spacious sagacious

shouldn’t shook should
boot booed boo
commission commodity commence communicate
appraisal atomizer Amazon scissors wizard
sheet sheep sheen
laconical lacrosse laborious laconic
look lookalike looking
Tom torpid tonsil toss
doom duplex doomed douchebag
sewage soothsayer soot
tad tap tag tat tally
dog dot dodge doctor
shoe shoot shoestring sure
hazard appetiser hazardous
tie tied tight aberrant
nil knitting Nick nibbles
abhorrence warren abracadabra
door Doris dogmatic dollar
Zaire zeppelin zealous Zen
loop loosen lunatic leukemia
kin kinship kissing killer
zonk zombie zigzag Zimbabwe
carbon calm cartwheel ratatouille
zoo zucchini zoom
knee niece needle neoclassical
dirt dirge dermatologist
soliloquy sabbatical satirical sarcastic
conductor commercial complexity commemorate
bought bored born bore
rapacious rapport who’d hod
water ladder letter fattest

keeper keen keyboard
goofy goose goulash guru
tawdry taunt taught talk
lid limp lisp lick
bangers hangar singer
terrier text temper tendril
gorse goggle Gough got
mother father nothing something
chord coarse cordial corner cork
Shaw shawl shop shopping
salute Samoa solicitor sufficient
carnage carpet car carbs
cad cattle catch canned
sarcasm sardine saga sample
god goblet gospel
land lad lack left
tune dune book good
dolphin dock domestic when
but bud buck bug
tuna tulip tutorial together
leather lemma Latin lexis
tooth tuba tool toucan
cask carcass castle calf
do dude doomsday duvet
sauna Saudi sock
cocky coffee cop commons
cupcake cusp cup comfort
ship shiver shun shovel
better batter Peter four
kick kindred kitchen kidnap
tummy tunnel tuft tug
gecko get ghetto guest
Tabasco tomato beetle centre
naughty nausea nautical nostalgia

bark barn path laugh
Nathan necessity napkin net
couple cover custard cut
guilt give gilded giggle
lapse legend less
head heard had
luxury London lovely lung
cooking cuckoo couldn't cushion
chagrin shellacking skirmisher
smasher fisher
ditch disk dish dizzy division
lurk learn lurch lurking
girth gurgle girdle
caustic caw caution choral
letter fatter ladder scatter
teaser tedious tepee tea
sea seed seek
tear teenager t-shirt
dense decoration dental definition
zinc zipper zip
sermon circa surfer surname
souvenir soup sousaphone
Gavin gap gadget guess
sad sandy Sam salt
sauce sorcerer sort
term turn terse turd
lasting lava larceny last
leader leaf league leakage
dart dark darling darkness
quiche keel key kiwi
lover luggage lump lumberjack
teach team teak teeth
tour pour sure bed
tub touch tundra tusk

gherkin girl Gertrud
geek guitar gear geeky
task tapas tarn
tension test technical telegraph
garden garbage gaga
summer substitute celebration
Gothic gorilla gossip
Taboo Tahiti tarantula Tahoma
Shanghai shell shepherd shambles
nudge nutmeg nonsuch nuts
Dan damp damn dad
deep deer deem deeper
hard had hud toothbrush
gimmick gift gig giddy
tackle taffy tax tan
look lookalike looking
solace sausage saw
end doll dull
galactic gavotte gazette
shock short shear share
shift shin shibboleth shindig
cooper cougar coo couscous
dip Dick dig discovery
debt deck decibel dedicate
bad back bag ban
noon noodle noose newspaper
berserk desert deserving
sequence seam seat season seize
neat negation needless Ninja
duplicate doing troubadour Hinduism
naan nasty narcotic sewer
hit hid hint put
dance loud lout cow
galore galosh gazebo gazelle

toddler tolerable top tongs Tonga
tardy target tar tart
Sergeant supper Sunday sucker
Durban dirty derby
nab nan nag native
accusor advertiser razor accusative
lot logic lobby lobster
continent conscience cost constable
lark lama laughing large

ten shed add bun
Darcy darn darkroom
noon nook nougat
nickname nicotine nifty nitwit
silk sulk gold take
nod nocturnal knob nominal
colonel curdle cursive curtain

Musical passages

Selected Exercises


46
p
52
mf
58
f
Play with different types of articulations: legato, tenuto, detached, staccato, & accented
63
mf
69
Pause for a full measure in between!
Use double-tongue for 16th notes.
75
mf
79
83
88
92
96
100

Etude

Hering

From Hering, 2011, p. 3.

Legato Etude

Crees/Gane

Reproduced with permission from Crees and Gane, 1988, p. 16.

Department of Linguistics
Matthias Heyne, Ph.D. student
Telephone: +64 3 364 2987 ext. 8131
Email: matthias.heyne@pg.canterbury.ac.nz

The influence of First Language on playing brass instruments:
Ultrasound recordings with brass players from different language groups

Information sheet for participants

Thank you for offering to participate in my research project on the influence of First Language on playing brass instruments. In my Ph.D. project at the University of Canterbury I am investigating whether one's First Language/s influence/s the way one plays a brass instrument. In this part of the study I am collecting exploratory data by using an ultrasound machine to document tongue positions and movements in speech and while playing the trombone.

I am myself a bass trombonist with degrees in orchestral music and jazz and for this reason I am interested in the question why national schools of playing seem to exist, while, at the same time, I have been trained in Linguistics as part of a teaching degree and hold an interest in the relationship between language and music.

Your involvement in this study will consist of (1) reading a wordlist in your native language and possibly another language you are very proficient in, (2) playing selected musical exercises using a pBone plastic trombone and a standard 6 1/2 AL mouthpiece provided by the researcher, and (3) filling in a short questionnaire about the languages you speak and how you learned to play the trombone. For parts (1) and (2) of the experiment, the positions and movements of your tongue will be recorded using an ultrasound machine with a specifically designed probe holder held in place below your chin; additionally an audio track will be recorded, as well as video of your face to allow observation of your embouchure and possible movements of the ultrasound probe holder. The maximum duration of the experiment shall not exceed two-and-a-half hours including setup and testing of the recording equipment, although it is expected to be shorter in most cases. Ultrasound imaging creates images of the tongue surface by scanning the soft tissue with an ultra-high-frequency sound wave that has no known *in vivo* bio effects and is a safe and non-invasive procedure.

Participation is voluntary and you have the right to withdraw at any stage without penalty. If you withdraw, I will remove information relating to you as long as it has not yet been published.

(continued on following page)

The results of this study may be published, but you may be assured of the complete confidentiality of the data gathered in this investigation: your identity will not be made public without your prior consent. To ensure confidentiality, the video and audio files of the recordings as well as any other files containing personal data will be labelled only with your participant number. Data that allows matching personal information with the recordings will be kept in a separate file and all files will be password-protected. Access to these files will be restricted to the researcher and others involved in the research project at the University of Canterbury, i.e. the supervisor and potential research associates.

The resulting Ph.D. thesis is a public document and will be available through the UC Library. You may receive a copy of the project results by contacting me at the conclusion of the project.

This project is being carried out as an exploratory study for my PhD research project at the University of Canterbury, Christchurch, New Zealand, under the supervision of Professor Jennifer Hay, who can be contacted at jen.hay@canterbury.ac.nz. Professor Hay will be pleased to discuss any concerns you may have about participation in the project.

This project has been reviewed and approved by the University of Canterbury Human Ethics Committee, and participants should address any complaints to The Chair, Human Ethics Committee, University of Canterbury, Private Bag 4800, Christchurch (human-ethics@canterbury.ac.nz).

If you agree to participate in the study, you are asked to complete the consent form and sign it before beginning this project.

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Department of Linguistics
Matthias Heyne, Ph.D. student
Telephone: +64 3 364 2987 ext. 8131
Email: matthias.heyne@pg.canterbury.ac.nz

The influence of First Language on playing brass instruments:
Ultrasound recordings with brass players from different language groups

Consent form

I have been given a full explanation of this project and have had the opportunity to ask questions.

I understand what is required of me if I agree to take part in the research.

I understand that participation is voluntary and I may withdraw at any time without penalty. Withdrawal of participation will also include the withdrawal of any information I have provided should this remain practically achievable.

I understand that any information or opinions I provide will be kept confidential to the researcher and his research associates and that any published or reported results will not identify the participants. I understand that the resulting thesis is a public document and will be available through the UC Library.

I understand that all data collected for the study will be kept in password protected electronic form.

I understand the risks associated with taking part and how they will be managed.

I understand that I am able to receive a report on the findings of the study by contacting the researcher at the conclusion of the project.

I understand that I can contact the researcher (contact information is provided above) or supervisor (Professor Jennifer Hay, jen.hay@canterbury.ac.nz) for further information. If I have any complaints, I can contact the Chair of the University of Canterbury Human Ethics Committee, Private Bag 4800, Christchurch (human-ethics@canterbury.ac.nz)

By signing below, I agree to participate in this research project.

name (please print)	date	signature

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

HUMAN ETHICS COMMITTEE

Secretary, Lynda Griffioen
Email: human-ethics@canterbury.ac.nz

Ref: HEC 2014/02/LR-PS

26 February 2014

Matthias Heyne
Linguistics Department
UNIVERSITY OF CANTERBURY

Dear Matthias

Thank you for forwarding to the Human Ethics Committee a copy of the low risk application you have recently made for your research proposal "The influence of first language on playing brass instruments exploratory study: ultrasound recordings with trombone players from different language groups".

I am pleased to advise that this application has been reviewed and I confirm support of the Department's approval for this project.

Please note that this approval is subject to the following:

- Please ensure that personal information is stored separately from the questionnaire.
- In the information sheet, please include the time involved for participants.

With best wishes for your project.

Yours sincerely

Lindsey MacDonald
Chair, Human Ethics Committee

University of Canterbury Private Bag 4800, Christchurch 8140, New Zealand. www.canterbury.ac.nz

E S

Betreff: publishing article in membership journal
Von: Matthias Heyne <matthias.heyne@pg.canterbury.ac.nz>
Datum: 07.08.2014 16:14
An: Human Ethics <human-ethics@canterbury.ac.nz>

Dear HEC Committee,

I have been offered to write a short article to appear in The Mouthpiece, a monthly publication distributed to the members of the Brass Band Association of New Zealand (http://www.brassbanz.org.nz/index.php/73-the-mouthpiece). This article will mainly serve to recruit participants for my online questionnaire (HEC APPLICATION 2013/39/LR-PS) and my exploratory ultrasound study (HEC 2014/02/LR-PS). While the first application didn't specify any way of recruiting participants, the wording was the following in the second application: The subjects will be recruited through my private connections in the music scene..."

Please let me know if the submission of the attached article (with possible small changes) would pose any problems regarding existing HEC approval.

Regards,
Matthias Heyne

Anhänge:

Matthias Heyne - Does the way we speak influence the way we play our brass instruments.pdf	45 Bytes

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

HUMAN ETHICS COMMITTEE

Secretary, Lynda Griffioen
Email: human-ethics@canterbury.ac.nz

Ref: HEC 2013/39/LR-PS & 2014/02/LR-PS

13 August 2014

Matthias Heyne
Department of Linguistics
UNIVERSITY OF CANTERBURY

Dear Matthias

Thank you for your request for an amendment to your research proposals "The influence of first language upon playing a brass instrument pilot study: online questionnaire" and "The influence of first language on playing brass instruments exploratory study: Ultrasound recordings with trombone players from different language groups" as outlined in your email dated 7 August 2014.

I am pleased to advise that this request has been considered and approved by the Human Ethics Committee.

Yours sincerely

Lindsey MacDonald
Chair, Human Ethics Committee

Betreff: HEC 2014/02/LR-PS
Von: Matthias Heyne <matthias.heyne@pg.canterbury.ac.nz>
Datum: 24.10.2014 15:09
An: Human Ethics <human-ethics@canterbury.ac.nz>
Kopie (CC): Jennifer Hay <jen.hay@canterbury.ac.nz>

Dear Human Ethics Committee,

I was granted Ethics clearance for my research using ultrasound imaging of the tongue in February. My application did not state that the recordings would have to take place in Christchurch and I'm writing to you to inform you that I'm going to travel to Tonga 15-27 November 2014 to record participants there. I am also considering travelling to Auckland with the next few months to record NZ English and/or Tongan participants and I recorded 4 participants in Germany in September 2014. The maximum number of subjects (25 as stated in my application) will not be exceeded by doing this.

On the Tonga trip I may also carry out research for Vica Papp of the NZILBB if she can get Ethics clearance before my trip commences.

Kind regards,
Matthias Heyne

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

HUMAN ETHICS COMMITTEE

Secretary, Lynda Griffioen
Email: human-ethics@canterbury.ac.nz

Ref: HEC 2014/02/LR-PS

3 November 2014

Matthias Heyne
Department of Linguistics
UNIVERSITY OF CANTERBURY

Dear Matthias

Thank you for your request for an amendment to your research proposal "The influence of first language on playing brass instruments exploratory study: ultrasound recordings with trombone players from different language groups" as outlined in your email dated 24 October 2014.

I am pleased to advise that this request has been considered and approved by the Human Ethics Committee.

Yours sincerely

Lindsey MacDonald
Chair, Human Ethics Committee

University of Canterbury Private Bag 4800, Christchurch 8140, New Zealand. www.canterbury.ac.nz

F E S

Matthias Heyne

From:	Matthias Heyne
Sent:	Monday, 8 December 2014 1:55 p.m.
To:	Human Ethics
Cc:	Jennifer Hay; Donald Derrick (donald.derrick@canterbury.ac.nz)
Subject:	HEC APPLICATION 2014/02/LR-PS - amendment to record further participants

Dear Human Ethics Committee,

I have now recorded 22 trombone players with different First Languages as specified in my application 2014/02/LR-PS. The study was initially planned as a pilot only but it has worked very well and doesn't need to be changed to collect more data for my PhD project so I am hereby applying for an amendment to let me record up to 20 additional players (a maximum of 45 in total).

Kind regards,
Matthias Heyne

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

HUMAN ETHICS COMMITTEE

Secretary, Lynda Griffioen
Email: human-ethics@canterbury.ac.nz

Ref: HEC 2014/02/LR-PS

8 December 2014

Matthias Heyne
Department of Linguistics
UNIVERSITY OF CANTERBURY

Dear Matthias

Thank you for your request for an amendment to your research proposal "The influence of first language on playing brass instruments exploratory study: ultrasound recordings with trombone players from different language groups" as outlined in your email dated 8 December 2014.

I am pleased to advise that this request has been considered and approved by the Human Ethics Committee.

Yours sincerely

Lindsey MacDonald
Chair, Human Ethics Committee

University of Canterbury Private Bag 4800, Christchurch 8140, New Zealand. www.canterbury.ac.nz

F E S

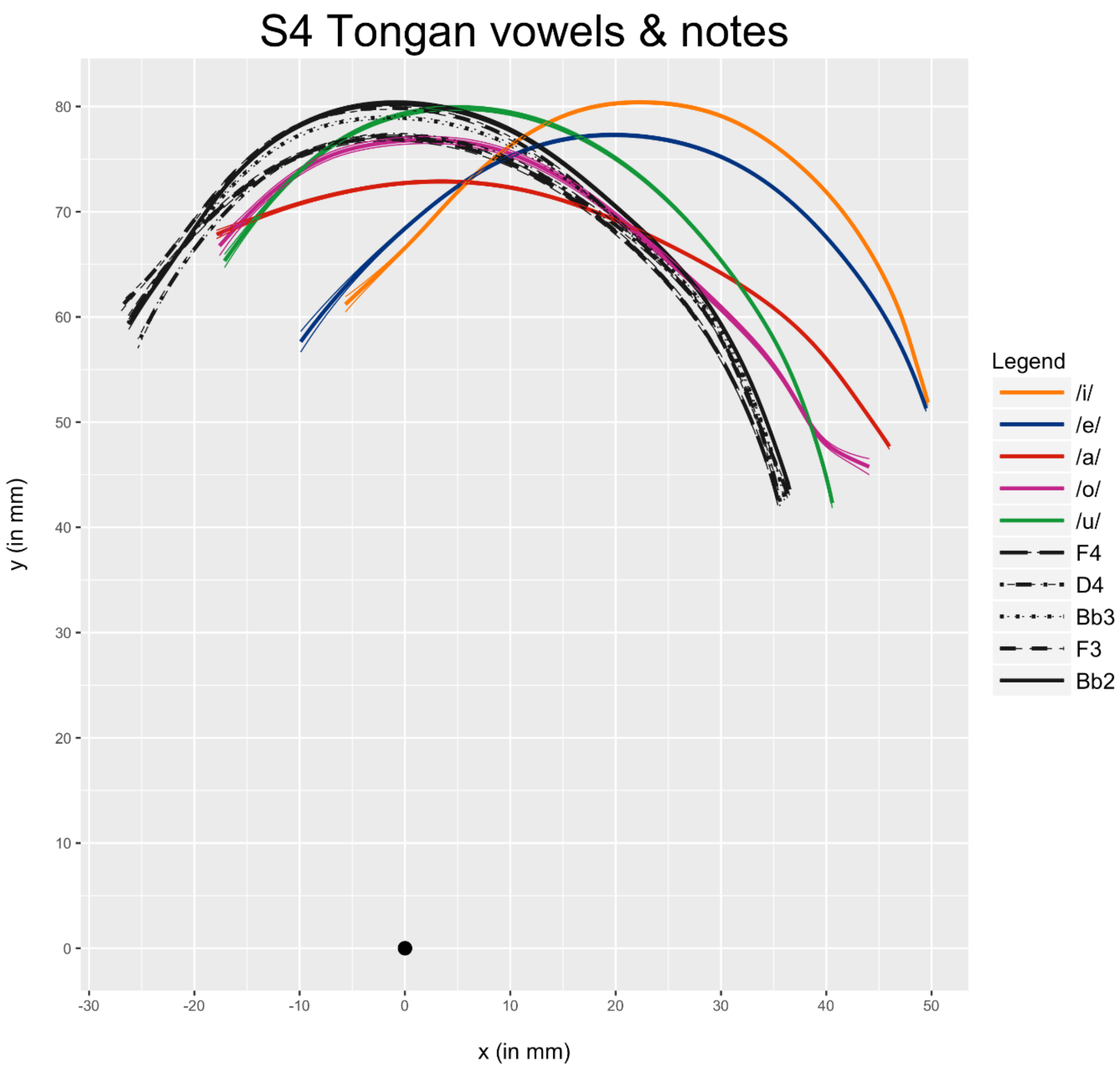
Appendix C: Z-scoring ratios and individual plots for all participants, including separate SSANOVAs for the five different notes analyzed for the musical passages

Z-scoring ratios for all participants

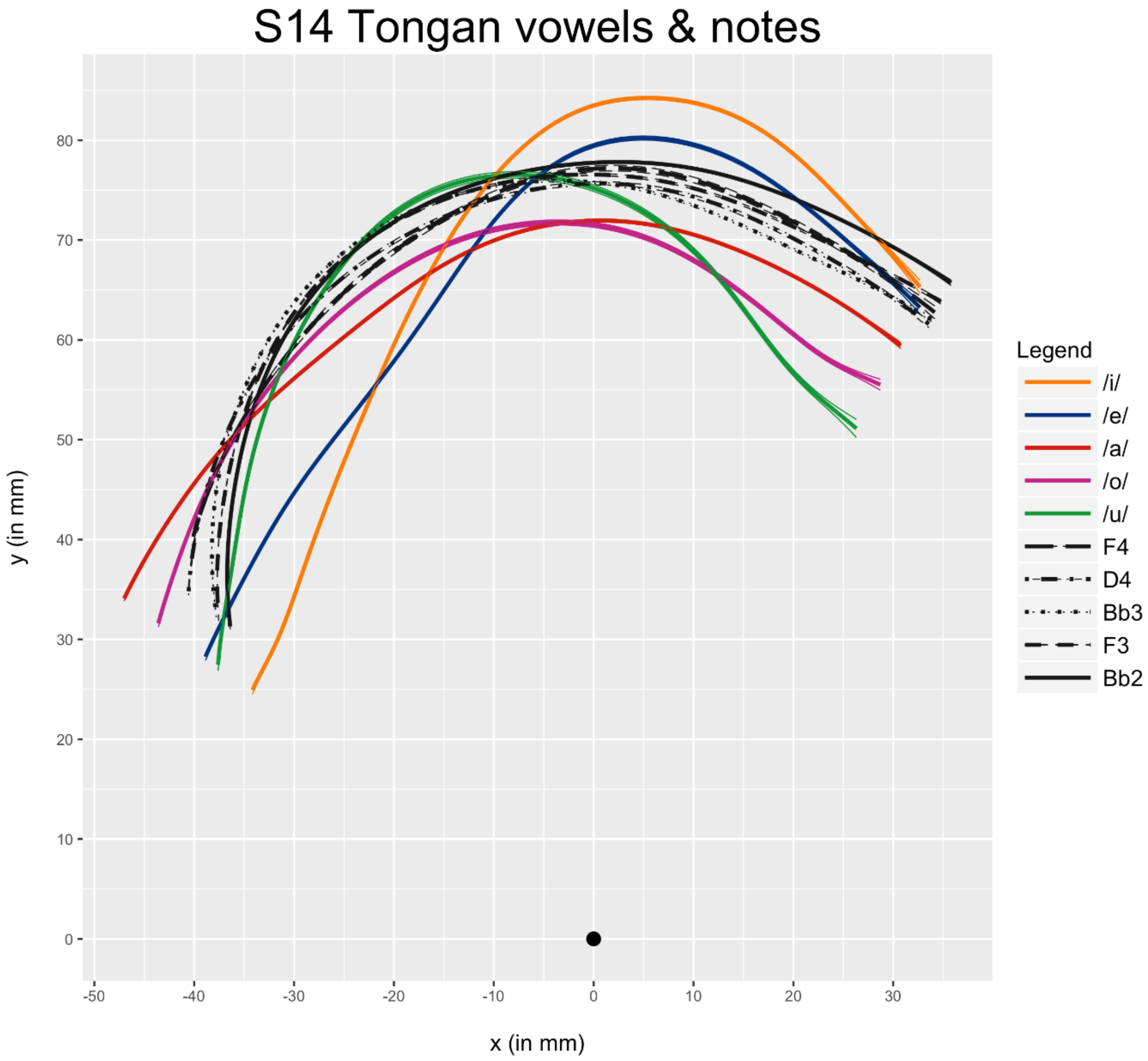
Tongan participants										
participant	S4	S14	S15	S16	S17	S18	S19	S20	S21	S22
Rotation (in degrees)	10.84	-5.49	-3.34	3.19	4.25	0.48	2.96	10.08	6.08	-0.81
z-scoring by distance measured for S24	1.07	0.97	0.89	0.97	0.90	0.97	0.96	0.95	0.93	0.94
NZE participants										
participant	S1	S3	S5	S12	S24	S25	S26	S27	S29	
Rotation (in degrees)	2.61	6.15	4.49	1.40	0.00	2.29	-5.07	-0.71	2.32	
z-scoring by distance measured for S24	0.69	0.92	0.96	0.98	1.00	0.98	1.06	1.06	1.12	

Plots for Tongan participants

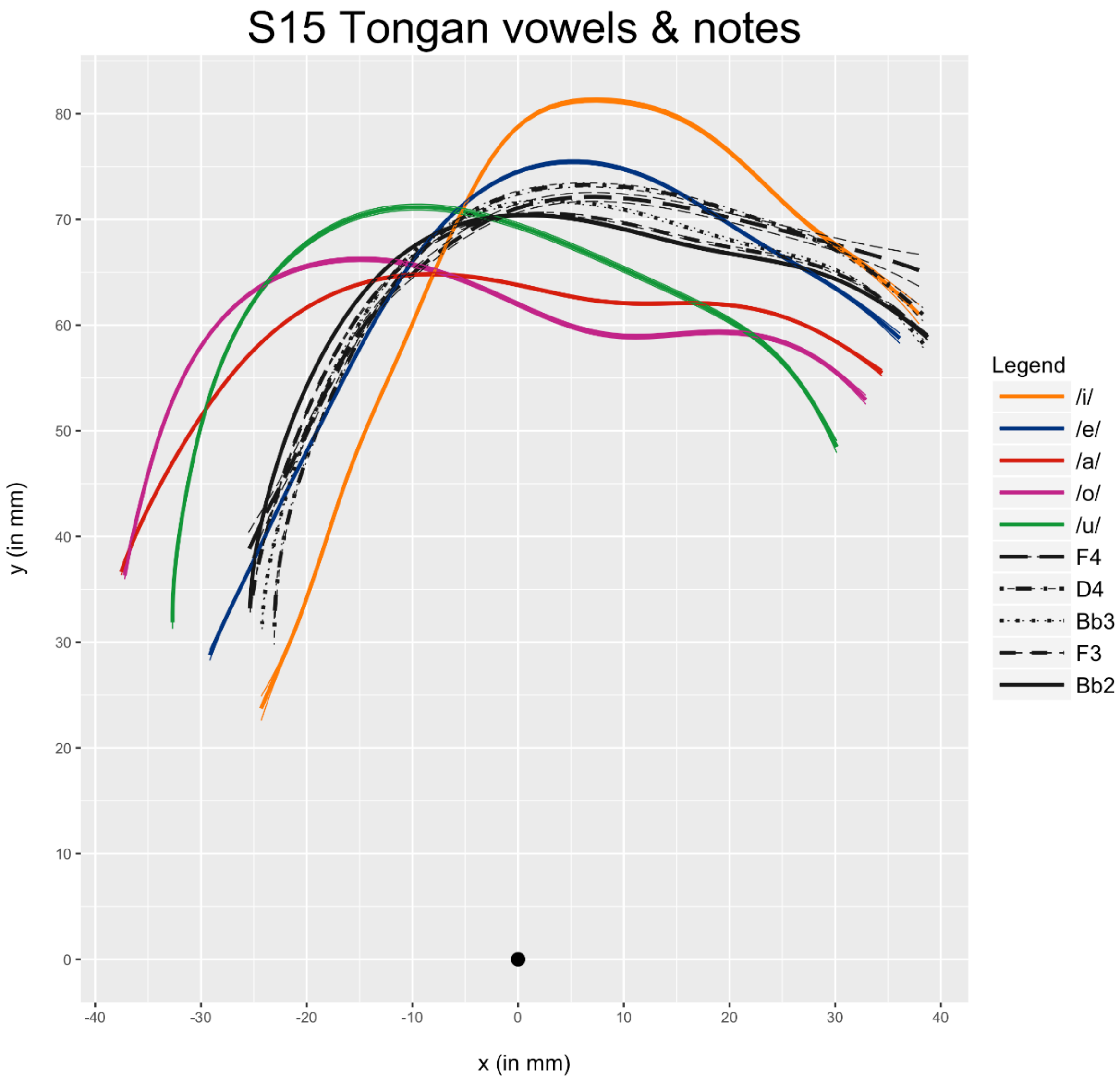
The following plots show the SSANOVA average midsagittal tongue curves for sustained note productions overlaid on the respective participant's vowel production in their native language, Tongan. The front of the tongue is to the right of the image while the back of the tongue is shown at the left. 95 percent confidence intervals are plotted as upper and lower bounds around the SSANOVA average curves (using the same colors), even though they are barely visible aside from at the edges. The scale for these plots is in mm and reflects the different sizes of the players' oral cavities (these data have not been rotated or normalized/z-scored).

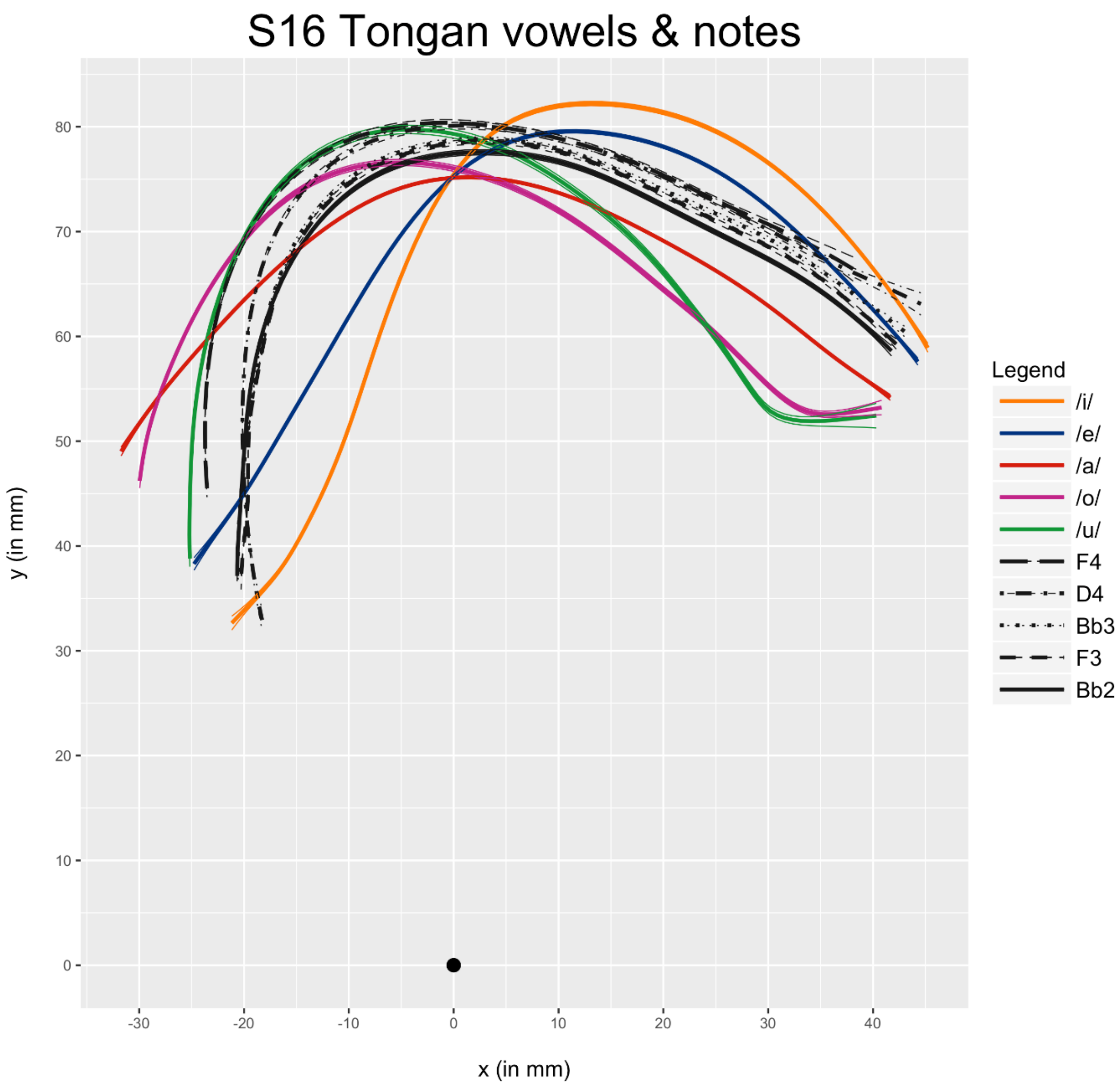


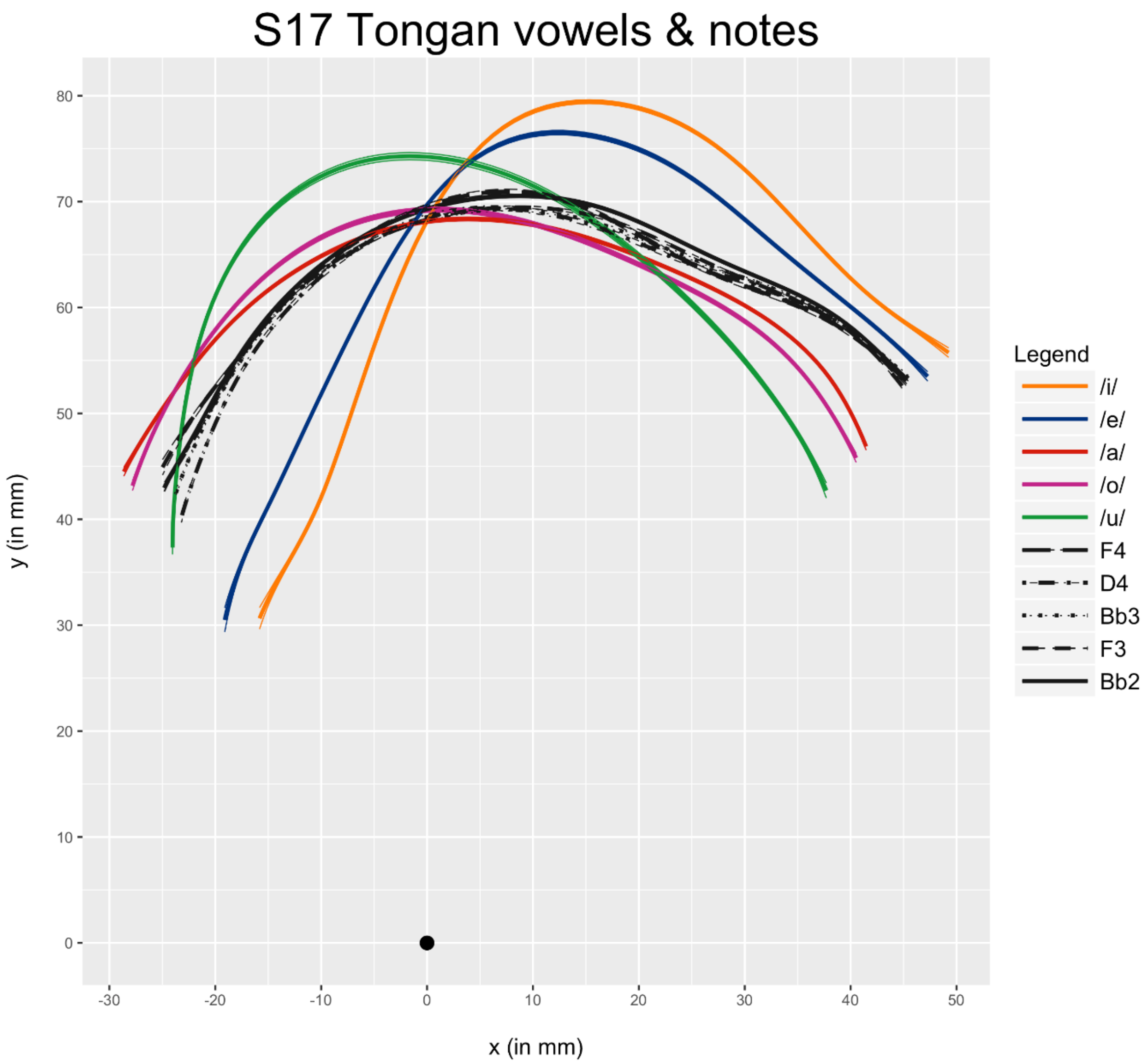
S4 Tongan vowels & notes
y (in mm)
x (in mm)
Legend
/i/
/e/
/a/
/o/
/u/
F4
D4
Bb3
F3
Bb2

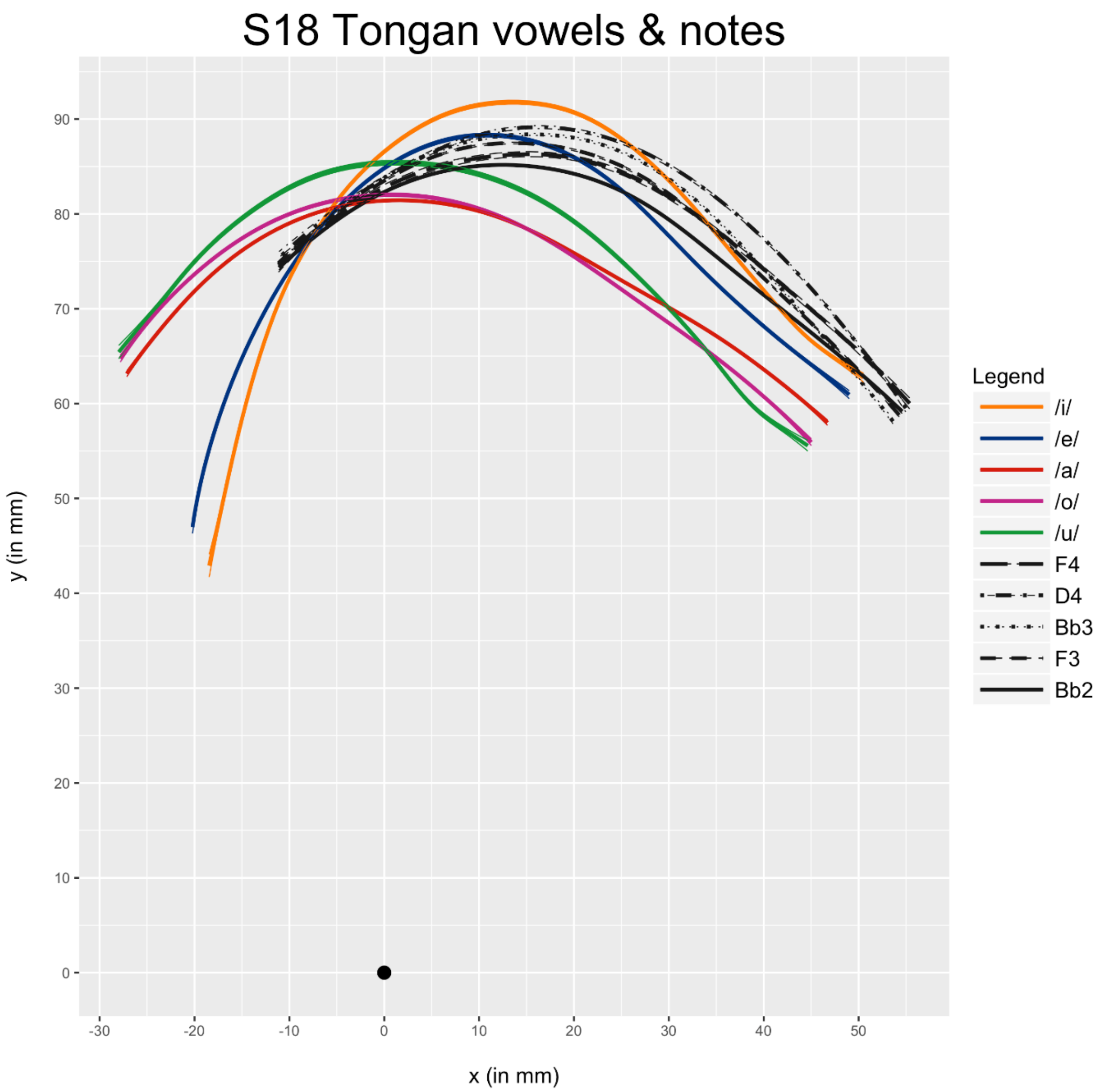


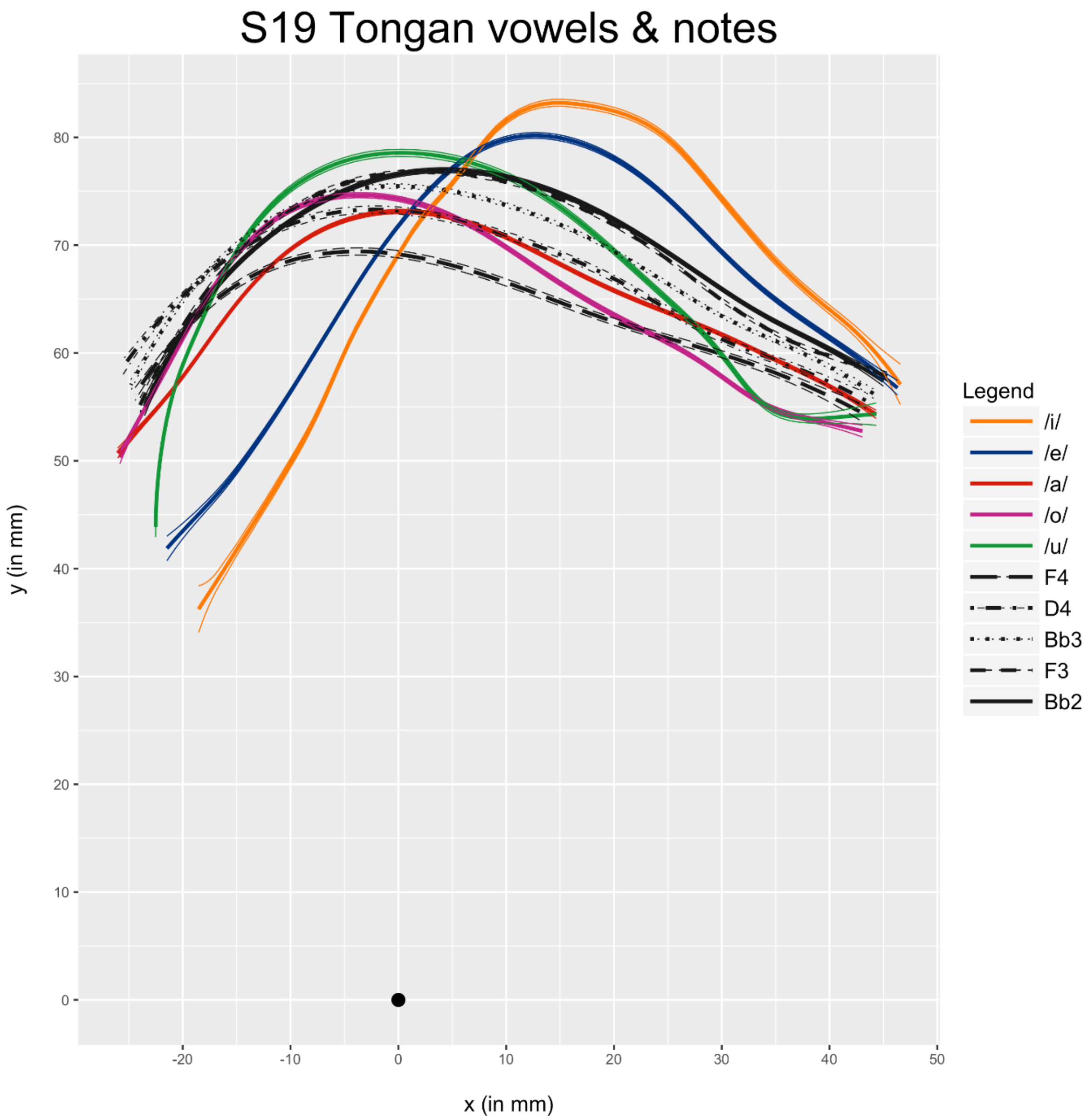
S14 Tongan vowels & notes
y (in mm)
x (in mm)
80
70
60
50
40
30
20
10
0
-50
-40
-30
-20
-10
0
10
20
30
Legend
/i/
/e/
/a/
/o/
/u/
F4
D4
Bb3
F3
Bb2


S15 Tongan vowels & notes
y (in mm)
x (in mm)
Legend
/i/
/e/
/a/
/o/
/u/
F4
D4
Bb3
F3
Bb2

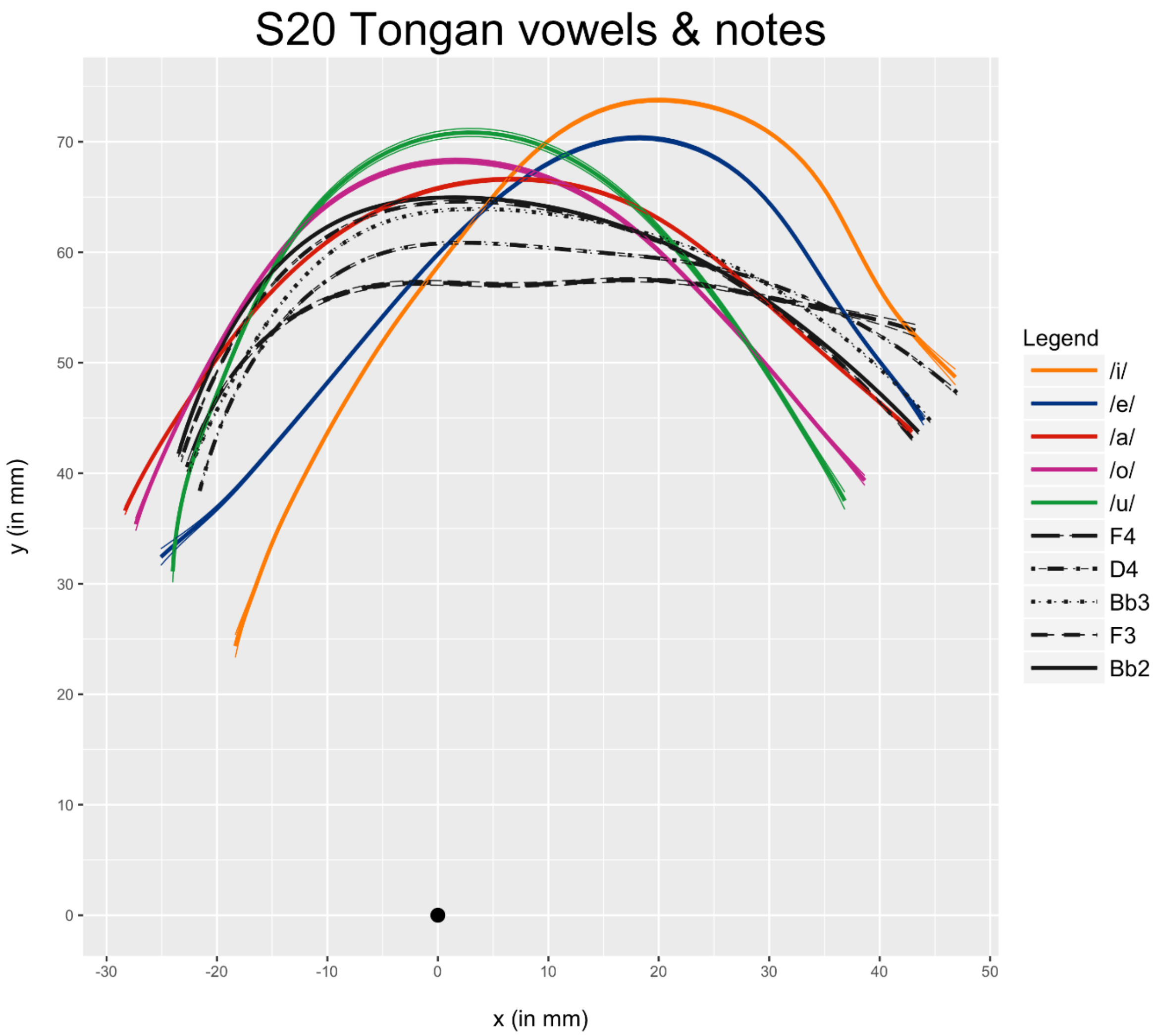

S16 Tongan vowels & notes
y (in mm)
x (in mm)
80
70
60
50
40
30
20
10
0
-30
-20
-10
0
10
20
30
40
Legend
/i/
/e/
/a/
/o/
/u/
F4
D4
Bb3
F3
Bb2

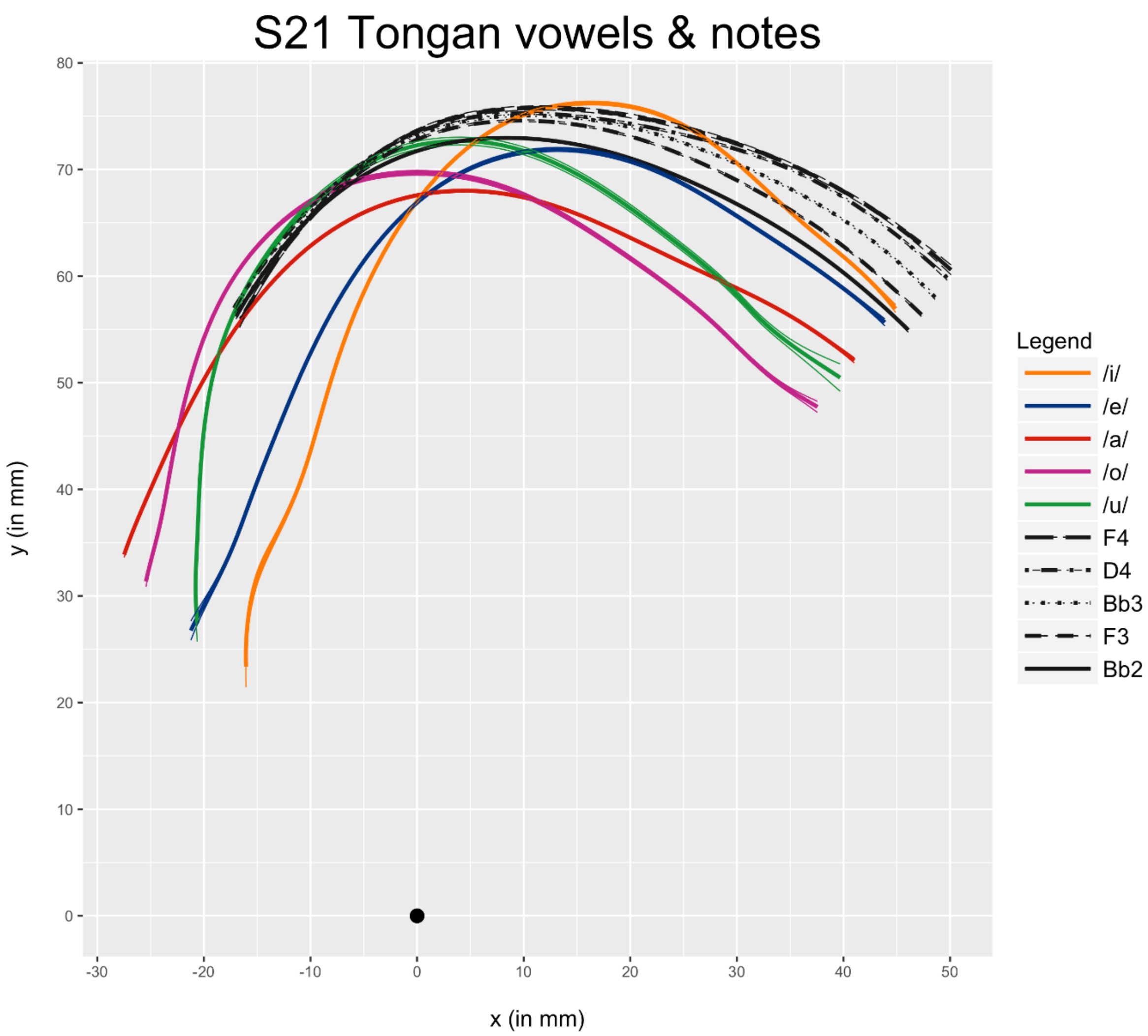

S17 Tongan vowels & notes
y (in mm)
x (in mm)
Legend
/i/
/e/
/a/
/o/
/u/
F4
D4
Bb3
F3
Bb2

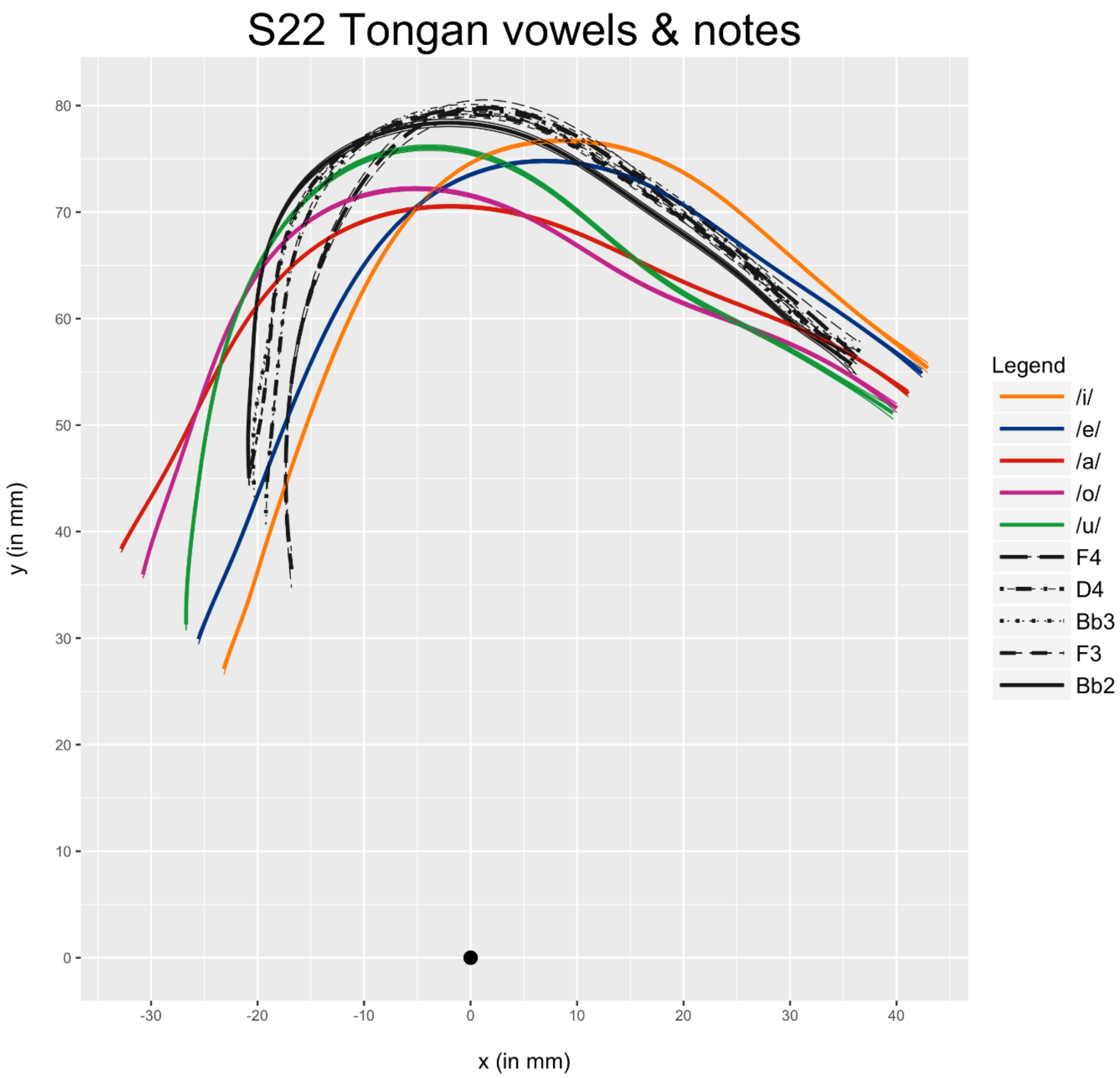

S18 Tongan vowels & notes
y (in mm)
x (in mm)
Legend
/i/
/e/
/a/
/o/
/u/
F4
D4
Bb3
F3
Bb2



S19 Tongan vowels & notes
y (in mm)
x (in mm)
Legend
/i/
/e/
/a/
/o/
/u/
F4
D4
Bb3
F3
Bb2


S20 Tongan vowels & notes
y (in mm)
x (in mm)
70
60
50
40
30
20
10
0
-30
-20
-10
0
10
20
30
40
50
Legend
/i/
/e/
/a/
/o/
/u/
F4
D4
Bb3
F3
Bb2

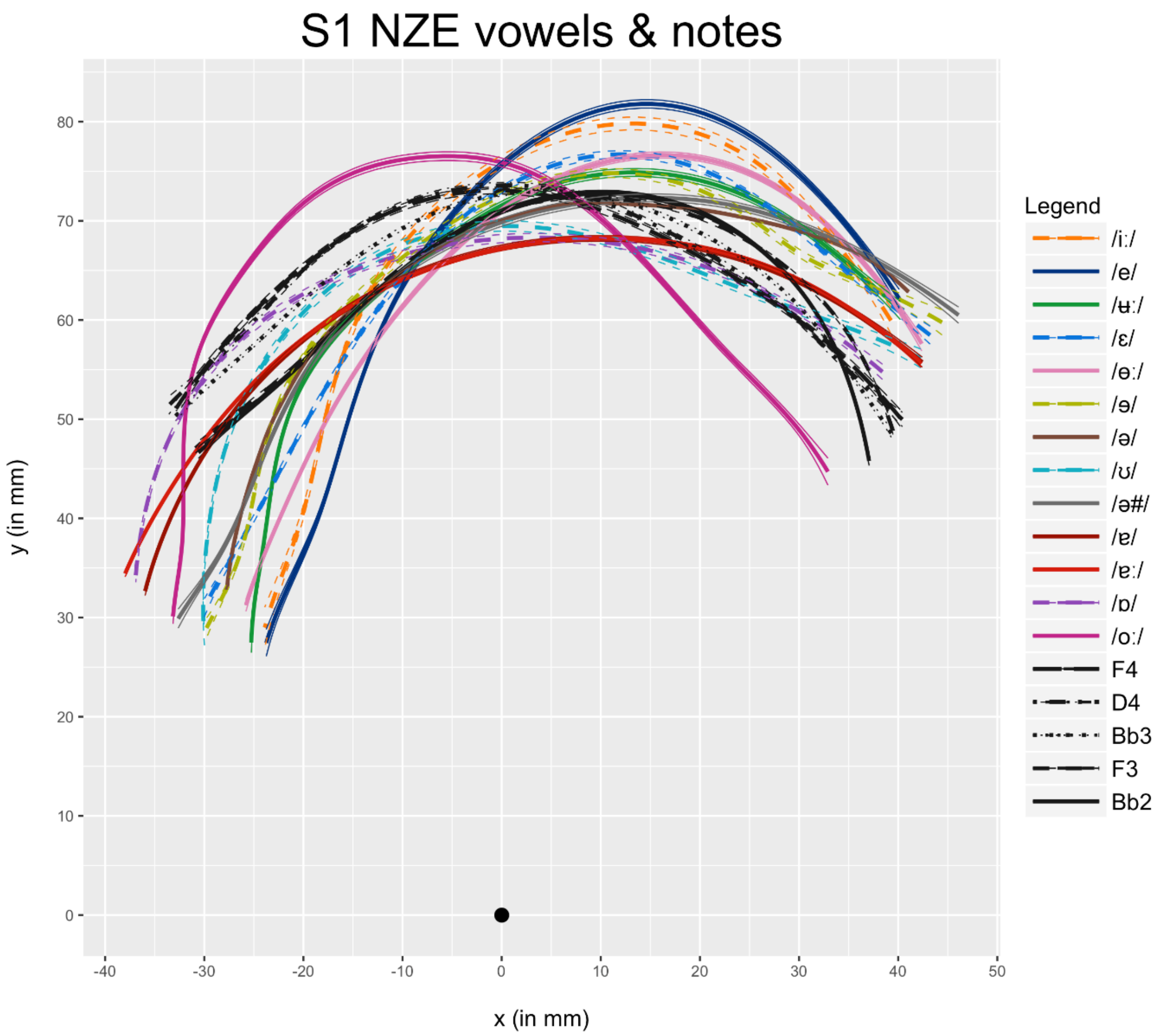

S21 Tongan vowels & notes
y (in mm)
x (in mm)
80
70
60
50
40
30
20
10
0
-30
-20
-10
0
10
20
30
40
50
Legend
/i/
/e/
/a/
/o/
/u/
F4
D4
Bb3
F3
Bb2



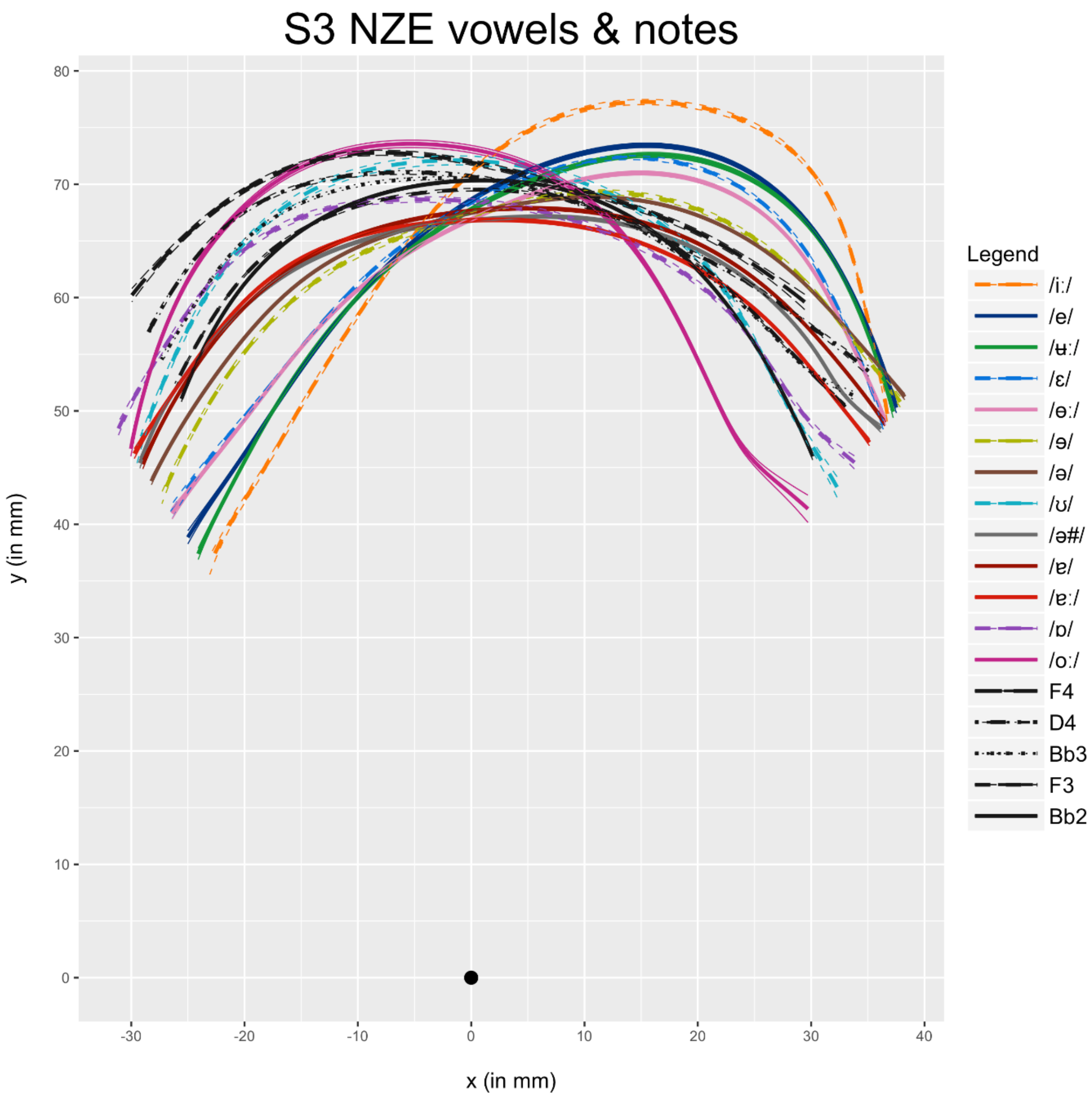
S22 Tongan vowels & notes
y (in mm)
x (in mm)
80
70
60
50
40
30
20
10
0
-30
-20
-10
0
10
20
30
40
Legend
/i/
/e/
/a/
/o/
/u/
F4
D4
Bb3
F3
Bb2

Plots for NZE participants

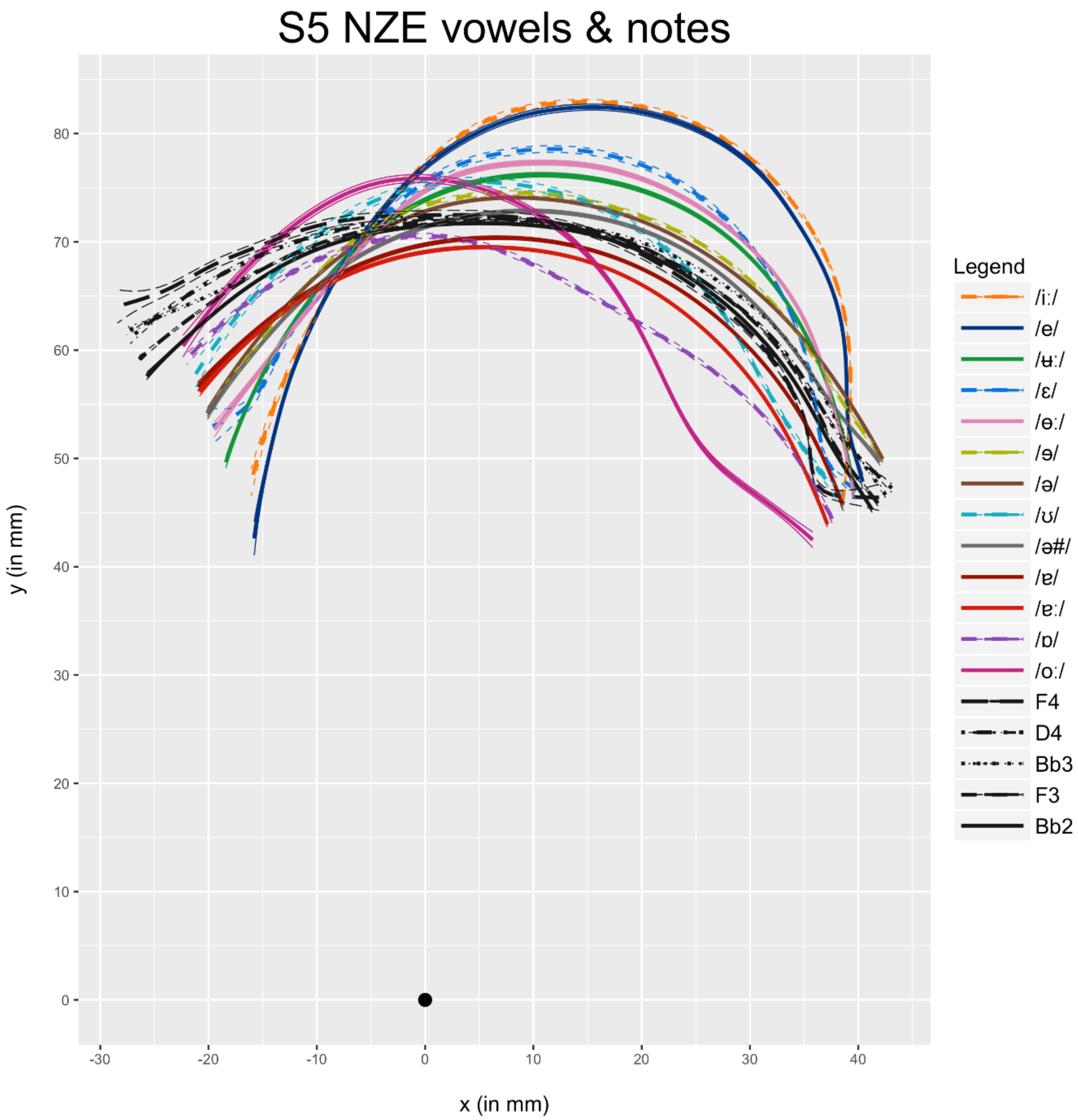
The following plots show the SSANOVA average midsagittal tongue curves for sustained note productions overlaid on the respective participant's vowel production in their native language, NZE. The front of the tongue is to the right of the image while the back of the tongue is shown at the left. 95 percent confidence intervals are plotted as upper and lower bounds around the SSANOVA average curves (using the same colors), even though they are barely visible aside from at the edges. The scale for these plots is in mm and reflects the different sizes of the players' oral cavities (these data have not been rotated or normalized/z-scored).

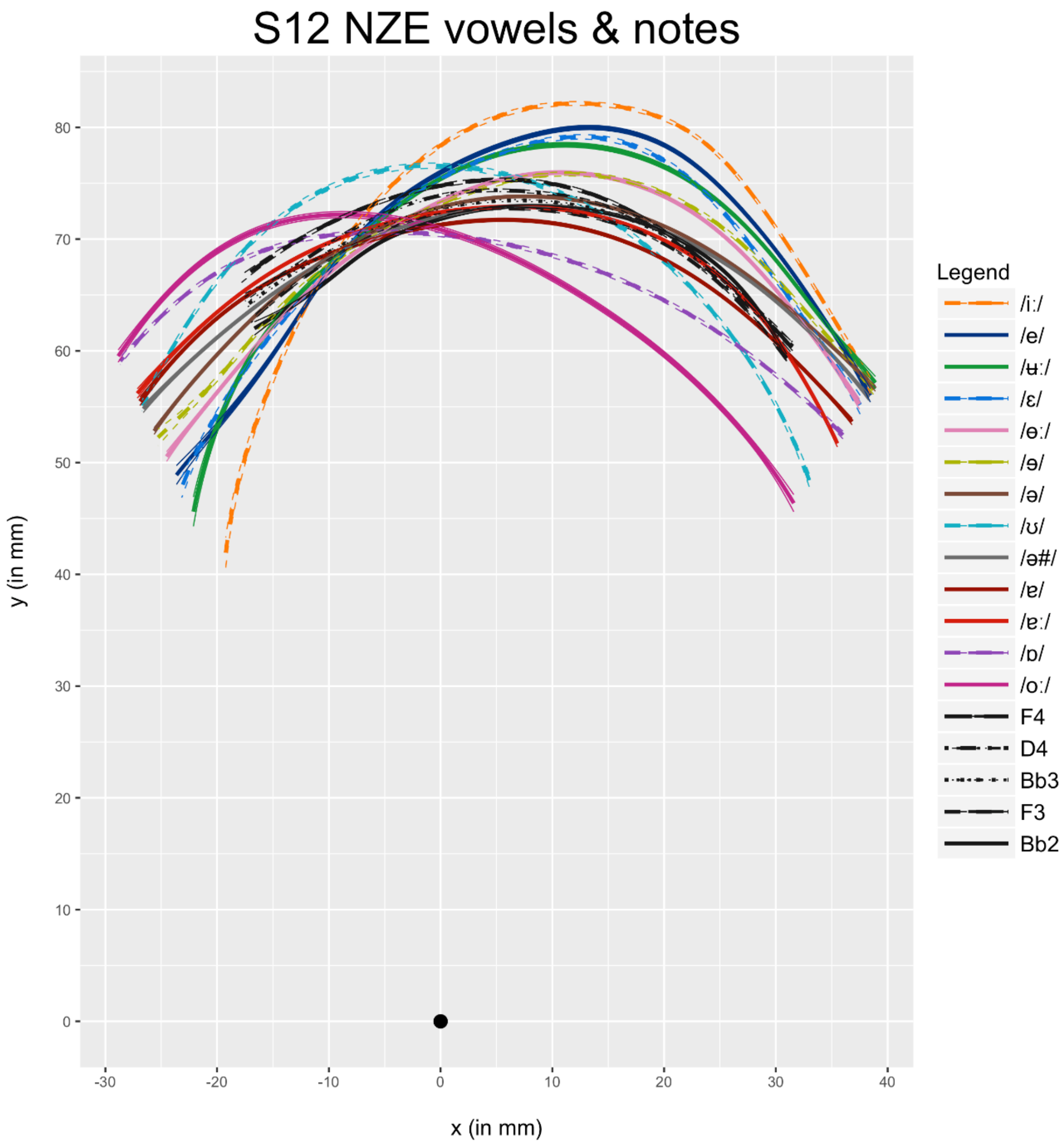


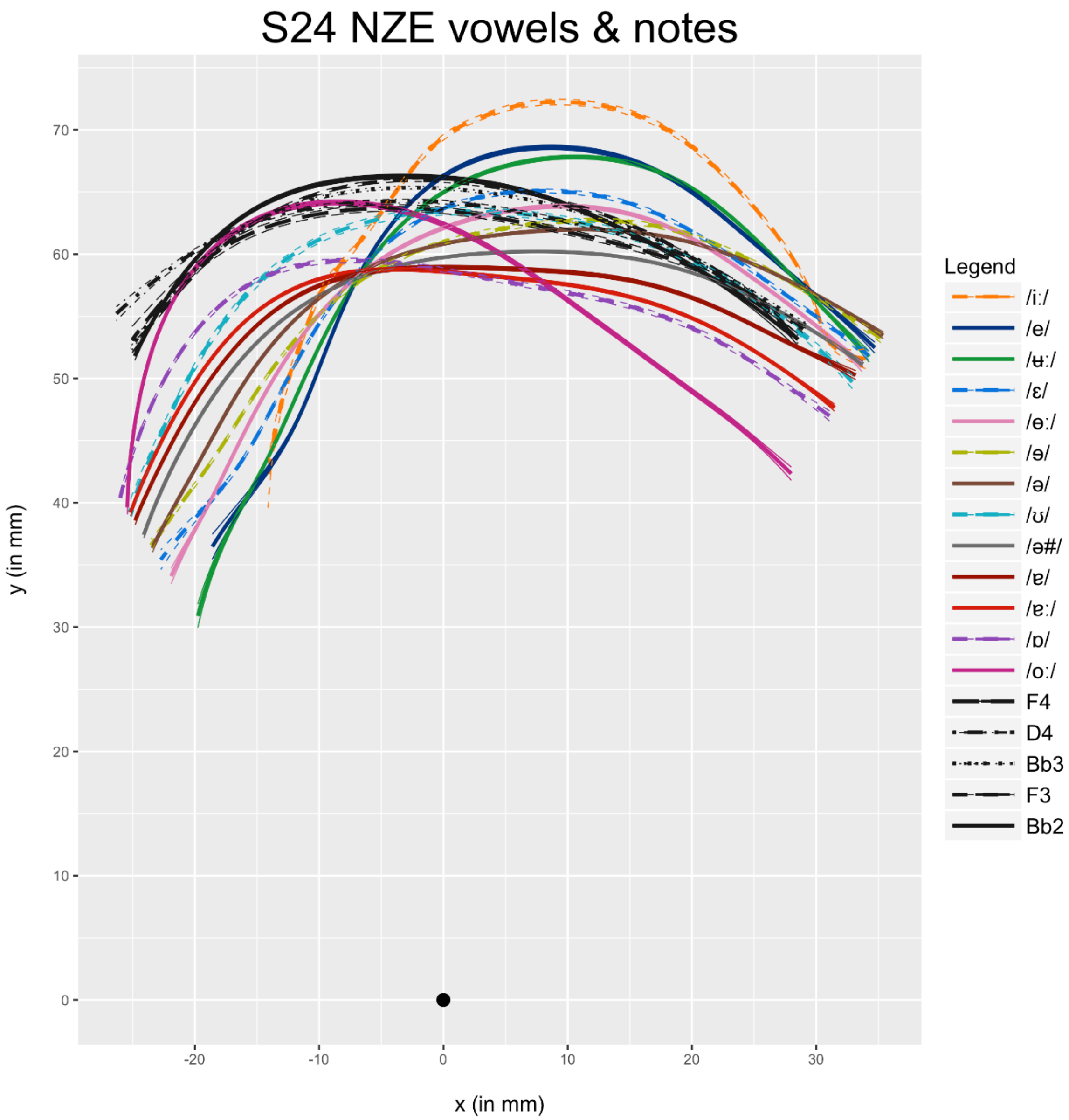
Note that S1 was the pilot participant and token numbers for vowel productions were much lower than for all other participants; this is reflected by the bigger SE bounds.

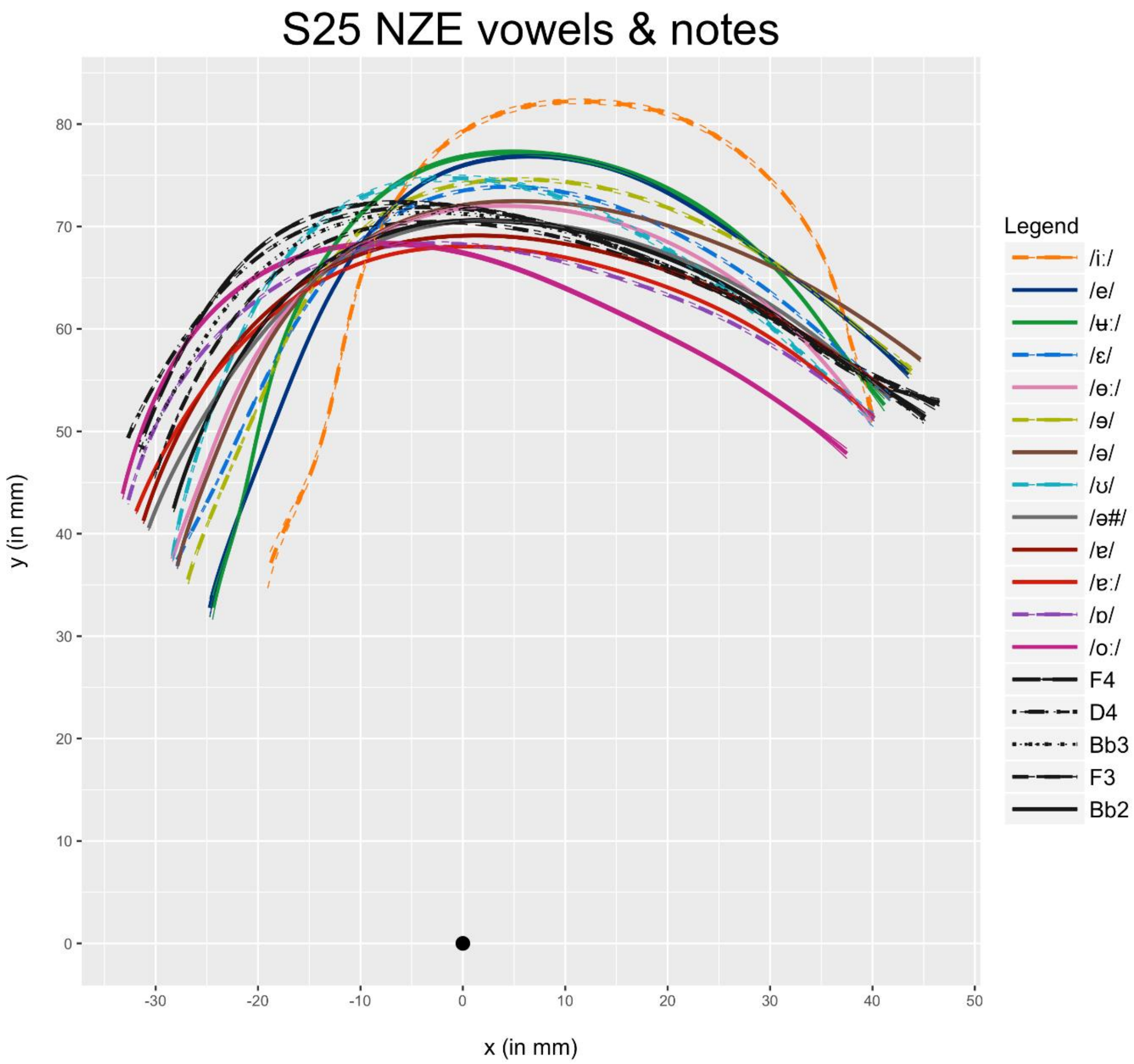


S3 NZE vowels & notes
y (in mm)
x (in mm)
Legend
/iː/
/e/
/ʉː/
/ɛ/
/ɵː/
/ɘ/
/ə/
/ʊ/
/ə#/
/ɐ/
/ɐː/
/ɒ/
/oː/
F4
D4
Bb3
F3
Bb2

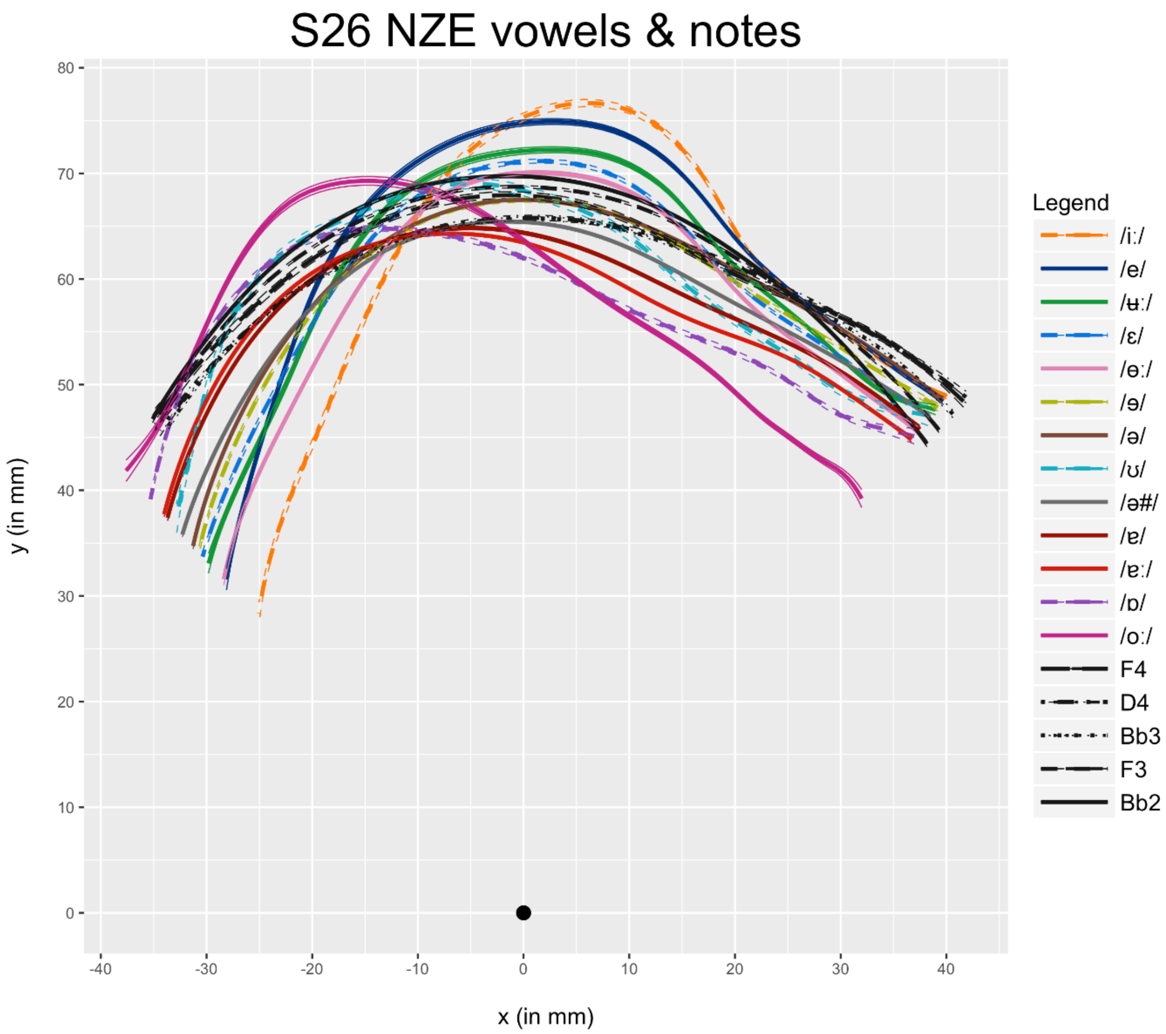

S5 NZE vowels & notes
y (in mm)
x (in mm)
Legend
/iː/
/e/
/ʉː/
/ɛ/
/eː/
/ɘ/
/ə/
/ʊ/
/ə#/
/ɐ/
/ɐː/
/ɒ/
/oː/
F4
D4
Bb3
F3
Bb2

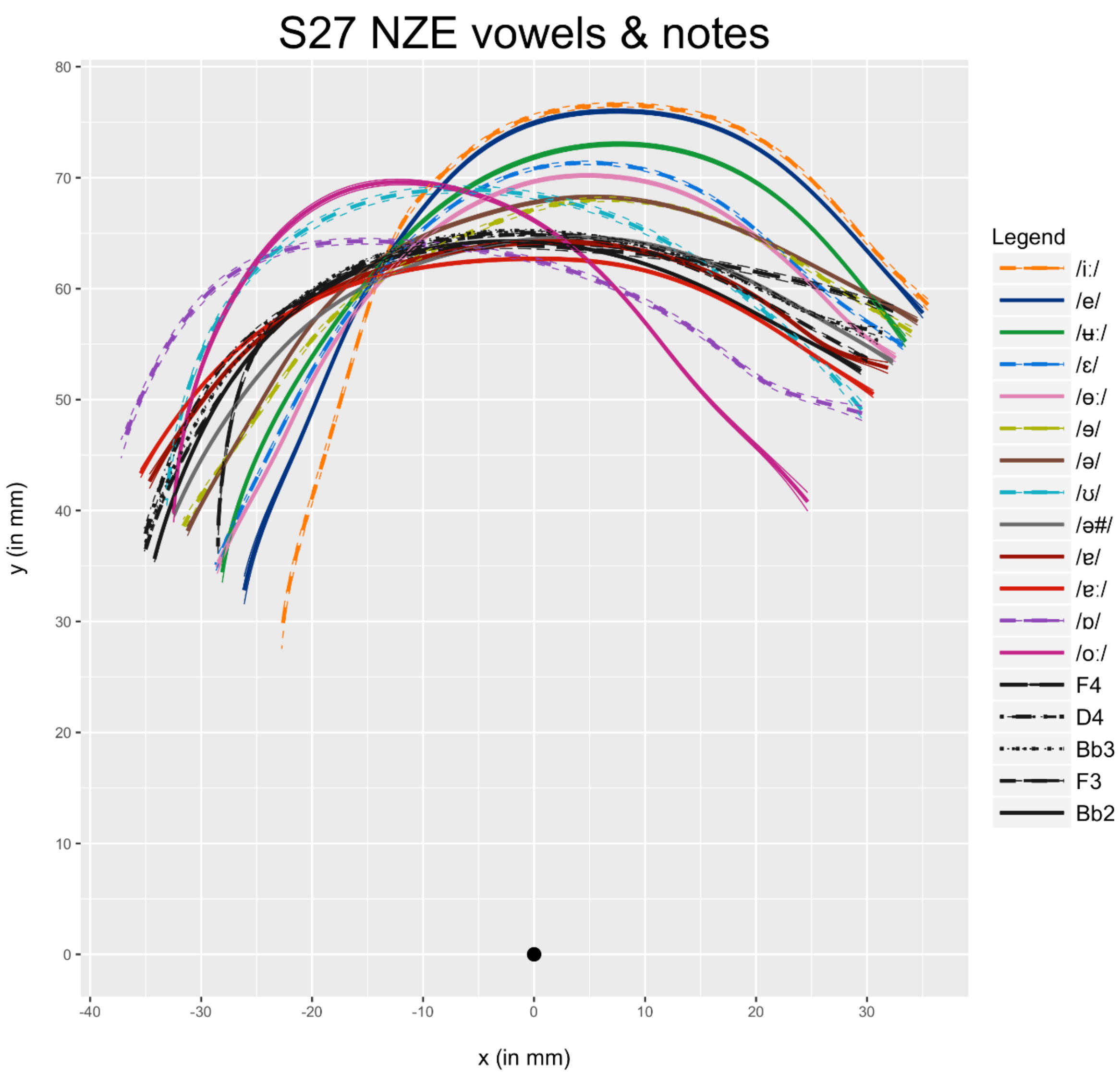

S12 NZE vowels & notes
y (in mm)
x (in mm)
80
70
60
50
40
30
20
10
0
-30
-20
-10
0
10
20
30
40
Legend
/iː/
/e/
/ʉː/
/ɛ/
/eː/
/ɘ/
/ə/
/ʊ/
/ə#/
/ɐ/
/ɐː/
/ɒ/
/oː/
F4
D4
Bb3
F3
Bb2


S24 NZE vowels & notes
y (in mm)
x (in mm)
Legend
/iː/
/e/
/ʉː/
/ɛ/
/ɵː/
/ɘ/
/ə/
/ʊ/
/ə#/
/ɐ/
/ɐː/
/ɒ/
/oː/
F4
D4
Bb3
F3
Bb2

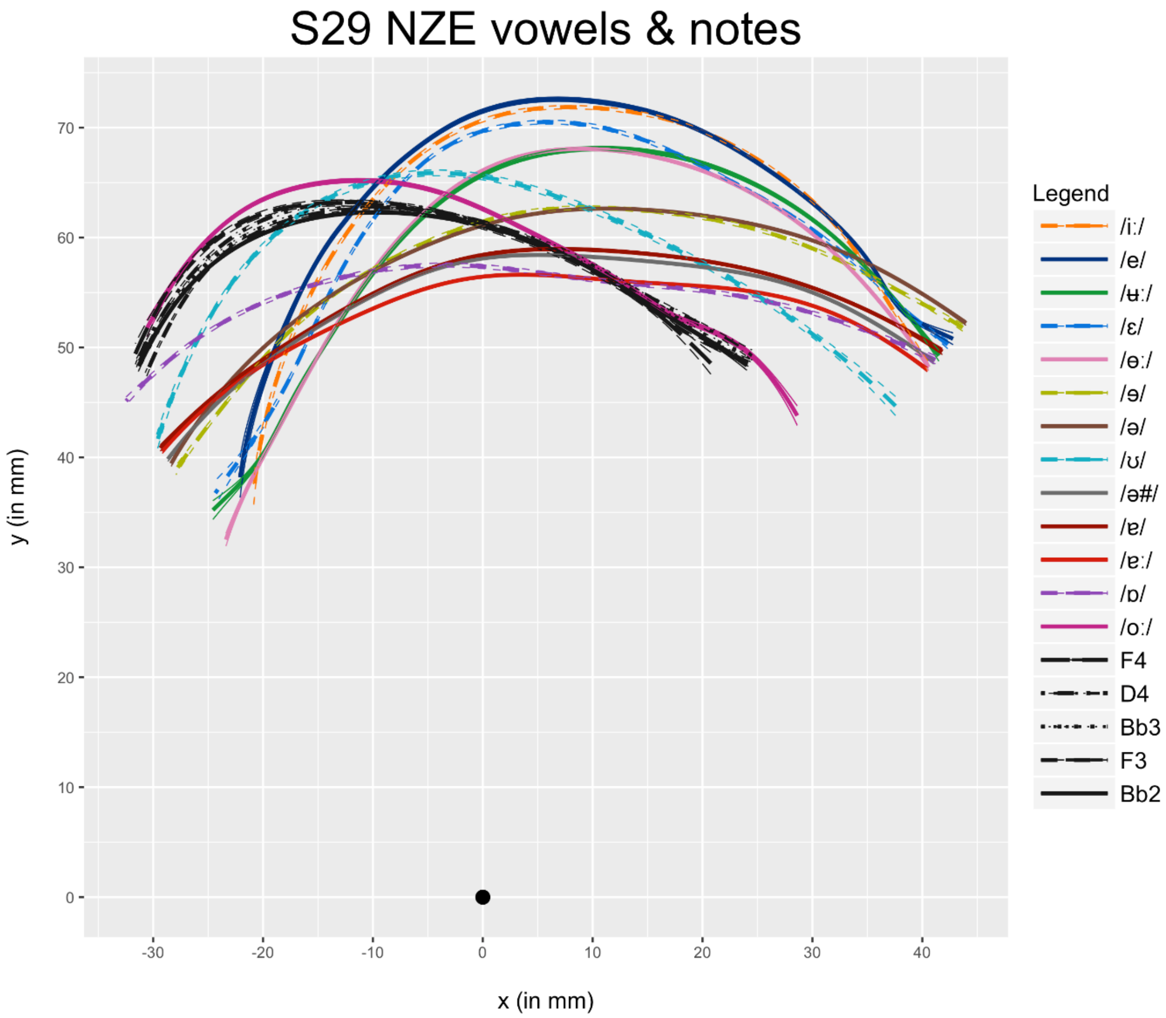


S25 NZE vowels & notes
Legend
/iː/
/e/
/ʉː/
/ɛ/
/eː/
/ɘ/
/ə/
/ʊ/
/ə#/
/ɐ/
/ɐː/
/ɒ/
/oː/
F4
D4
Bb3
F3
Bb2
y (in mm)
x (in mm)


S26 NZE vowels & notes
y (in mm)
x (in mm)
Legend
/iː/
/e/
/ʉː/
/ɛ/
/ɵː/
/ɘ/
/ə/
/ʊ/
/ə#/
/ɐ/
/ɐː/
/ɒ/
/oː/
F4
D4
Bb3
F3
Bb2



S27 NZE vowels & notes
80
70
60
50
40
30
20
10
0
-40
-30
-20
-10
0
10
20
30
y (in mm)
x (in mm)
Legend
/iː/
/e/
/ʉː/
/ɛ/
/ɵː/
/ɘ/
/ə/
/ʊ/
/ə#/
/ɐ/
/ɐː/
/ɒ/
/oː/
F4
D4
Bb3
F3
Bb2



S29 NZE vowels & notes
y (in mm)
x (in mm)
70
60
50
40
30
20
10
0
-30
-20
-10
0
10
20
30
40
Legend
/iː/
/e/
/ʉː/
/ɛ/
/ɵː/
/ɘ/
/ə/
/ʊ/
/ə#/
/ɐ/
/ɐː/
/ɒ/
/oː/
F4
D4
Bb3
F3
Bb2

Appendix D: Permissions from publishers

Email from Peter, publisher of the Journal of the International Trumpet Guild

Subject: Re: ITG Journal article

From: Peter Wood <editor@trumpetguild.org>

Date: 25-Aug-16 11:26

To: Matthias Heyne <ma$hias.heyne@pg.canterbury.ac.nz>

Dear Matthias:

You hereby have permission to include your article:

Heyne, M., & Derrick, D. (2016). Visualization techniques for empirical brass instrument research. Journal of the International Trumpet Guild, 40, 6-14, 24

as part of your dissertation.

Please include the following statement: "The International Trumpet Guild grants permission to post this article in this format. For more information on ITG, visit their website (www.trumpetguild.org)."

Sincerely,

Peter

--

Dr. Peter Wood

ITG Publications Editor

(251) 533-1208